



# Central Limit Theorem for the Multilevel Monte Carlo Euler Method and Applications to Asian Options

Mohamed Ben Alaya, Ahmed Kebaier

## ► To cite this version:

Mohamed Ben Alaya, Ahmed Kebaier. Central Limit Theorem for the Multilevel Monte Carlo Euler Method and Applications to Asian Options. 2012. hal-00693191

**HAL Id: hal-00693191**

**<https://hal.science/hal-00693191>**

Preprint submitted on 2 May 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# CENTRAL LIMIT THEOREM FOR THE MULTILEVEL MONTE CARLO EULER METHOD AND APPLICATIONS TO ASIAN OPTIONS

MOHAMED BEN ALAYA<sup>1</sup> & AHMED KEBAIER<sup>\*1</sup>

<sup>1</sup>LAGA, CNRS (UMR 7539), Université Paris 13,  
99, av. J.B. Clément 93430 Villetaneuse, France  
mba@math.univ-paris13.fr kebaier@math.univ-paris13.fr

May 2, 2012

## Abstract

This paper focuses on studying the multilevel Monte Carlo method recently introduced by Giles [8] and significantly more efficient than the classical Monte Carlo one. Our aim is to prove a central limit theorem of Lindeberg Feller type for the multilevel Monte Carlo method associated to the Euler discretization scheme. To do so, we prove first a stable law convergence theorem, in the spirit of Jacod and Protter [15], for the Euler scheme error on two consecutive levels of the algorithm. This leads to an accurate description of the optimal choice of parameters and to an explicit characterization of the limiting variance in the central limit theorem of the algorithm. We investigate the application of the Multilevel Monte Carlo method to the pricing of Asian options, by discretizing the integral of the payoff process using Riemann and trapezoidal schemes. In particular, we prove stable law convergence for the error of these second order schemes. This allows us to prove two additional central limit theorems providing us the optimal choice of the parameters with an explicit representation of the limiting variance. For this setting of second order schemes, we give new optimal parameters leading to the convergence of the central limit theorem. Complexity analysis of the Multilevel Monte Carlo algorithm were processed.

**AMS 2000 Mathematics Subject Classification.** 60F05, 62F12, 65C05, 60H35.

**Key Words and Phrases.** Central limit theorem, Multilevel Monte Carlo methods, Euler scheme, Asian options, finance.

---

<sup>\*</sup>This research benefited from the support of the chair "Risques Financiers", Fondation du Risque.

# 1 Introduction

In many applications, in particular for the pricing of financial securities, we are interested in the effective computation by Monte Carlo methods of the quantity  $\mathbb{E}f(X_T)$ , where  $X := (X_t)_{0 \leq t \leq T}$  is a diffusion process and  $f$  a given function. The Monte Carlo Euler method consists of two steps. First, approximate the diffusion process  $(X_t)_{0 \leq t \leq T}$  by the Euler scheme  $(X_t^n)_{0 \leq t \leq T}$  with time step  $T/n$ . Then, approximate  $\mathbb{E}f(X_T^n)$  by  $\frac{1}{N} \sum_{i=1}^N f(X_{T,i}^n)$ , where  $f(X_{T,i}^n)_{1 \leq i \leq N}$  is a sample of  $N$  independent copies of  $f(X_T^n)$ . This approximation is affected respectively by a discretization error and a statistical error

$$\varepsilon_n := \mathbb{E}(f(X_T^n) - f(X_T)) \quad \text{and} \quad \frac{1}{N} \sum_{i=1}^N f(X_{T,i}^n) - \mathbb{E}f(X_T^n).$$

The optimal choice of the sample size  $N$  in the classical Monte Carlo method mainly depends on the order of the discretization error. In the context of possibly degenerate diffusions  $X$  and  $\mathcal{C}^1$ -functions  $f$ , Kebaier [16] proves that the rate of convergence of the discretization error  $\varepsilon_n$  can be  $1/n^\alpha$  for all values of  $\alpha \in [1/2, 1]$ . It turns out that for such order of convergence the optimal choice of  $N$  is given by  $n^{2\alpha}$ . This leads to a total complexity in the Monte Carlo method of order  $C_{MC} = n^{2\alpha+1}$ . A further discussion of this is given in subsection 2.4 (see Duffie and Glynn [5] for related results).

In order to improve the performance of this method, Kebaier [16] introduced a two-level Monte Carlo method [16] (called the statistical Romberg method) reducing the complexity  $C_{MC}$  while maintaining the convergence of the algorithm. This method uses two Euler schemes with time steps  $T/n$  and  $T/n^\beta$ ,  $\beta \in (0, 1)$  and approximates  $\mathbb{E}f(X_T)$  by

$$\frac{1}{N_1} \sum_{i=1}^{N_1} f(\hat{X}_{T,i}^{n^\beta}) + \frac{1}{N_2} \sum_{i=1}^{N_2} f(X_{T,i}^n) - f(X_{T,i}^{n^\beta}),$$

where  $\hat{X}_T^{n^\beta}$  is a second Euler scheme with time step  $T/n^\beta$  and such that the Brownian paths used for  $X_T^n$  and  $\hat{X}_T^{n^\beta}$  has to be independent of the Brownian paths used to simulate  $\hat{X}_T^{n^\beta}$ . In order to get a rational choice of  $N_1$ ,  $N_2$  and  $\beta$  versus  $n$ , Kebaier [16] proves a Central Limit Theorem for this new algorithm. This theorem uses the weak convergence of the normalized error of the Euler scheme for diffusions proved by Kurtz and Protter [19] (and strengthened by Jacod and Protter [15]). It turns out that for a given discretization error  $\varepsilon_n = 1/n^\alpha$  ( $\alpha \in [1/2, 1]$ ), the optimal choice is obtained for  $\beta = 1/2$ ,  $N_1 = n^{2\alpha}$  and  $N_2 = n^{2\alpha-(1/2)}$ . With this choice, the complexity of the statistical Romberg method is of order  $C_{SR} = n^{2\alpha+(1/2)}$  which is lower than the classical complexity in the Monte Carlo method.

More recently, Giles [8] generalized the statistical Romberg method of Kebaier [16] and proposed the multilevel Monte Carlo algorithm, in a similar approach to Heinrich's multilevel method for parametric integration [12] (see also Creutzig, Dereich, Müller-Gronbach and Ritter [3], Dereich [4], Giles [7], Giles, Higham and Mao [9], Heinrich [11] and Heinrich and Sindambiwe [13] for related results). The multilevel Monte Carlo method uses information from a sequence of computations with decreasing step sizes and approximates the quantity  $\mathbb{E}f(X_T)$  by

$$Q_n = \frac{1}{N_0} \sum_{k=1}^{N_0} f(X_{T,k}^1) + \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left( f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) \right), \quad m \in \mathbb{N} \setminus \{0, 1\} \quad \text{and} \quad L = \frac{\log n}{\log m}.$$

The process  $(X_t^{m^\ell})_{0 \leq t \leq T}$  denotes the Euler scheme with time step  $m^{-\ell}T$  for  $\ell \in \{0, \dots, L\}$ . Here, it is important to point out that all these  $L+1$  Monte Carlo estimators have to be based on different, independent samples. However, for fixed  $k$  and  $\ell$ , the simulations  $f(X_{T,k}^{m^\ell})$  and  $f(X_{T,k}^{m^{\ell-1}})$  have to be based in the same Brownian path but with different times steps  $m^{-\ell}T$  and  $m^{-(\ell-1)}T$ . Due to the above independence assumption for the paths, the variance of the multilevel estimator is given by

$$\sigma^2 := \text{Var}(Q_n) = N_0^{-1} \text{Var}(f(X_T^1)) + \sum_{\ell=1}^L N_\ell^{-1} \sigma_\ell^2,$$

where  $\sigma_\ell^2 = \text{Var}(f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}))$ . For a Lipschitz continuous function  $f$ , it is easy to check, using properties of the Euler scheme, that

$$\sigma^2 \leq c_2 \sum_{\ell=0}^L N_\ell^{-1} m^{-\ell}$$

for some positive constant  $c_2$  (see Proposition 1 for more details). Giles [8] uses this computation in order to find the optimal choice of the multilevel Monte Carlo parameters. More precisely, to obtain a desired root mean squared error (RMSE), say of order  $1/n^\alpha$ , for his multilevel estimator, Giles [8] uses the above computation on  $\sigma^2$  to minimize the total complexity of the algorithm. It turns out that the optimal choice is obtained for (see Theorem 3.1 of [8])

$$N_\ell = 2c_2 n^{2\alpha} \left( \frac{\log n}{\log m} + 1 \right) \frac{T}{m^\ell}, \quad \text{for } \ell \in \{0, \dots, L\} \quad \text{and} \quad L = \frac{\log n}{\log m}. \quad (1)$$

This optimal choice leads to a complexity for the multilevel Monte Carlo Euler method proportional to  $n^{2\alpha}(\log n)^2$ . Interesting numerical tests, comparing three methods (crude Monte Carlo, Statistical Romberg and the Multilevel Monte Carlo), were proceed in Korn, Korn and Kroisandt [18]. Furthermore, Giles [8] obtain also the optimal parameters for the multilevel Monte Carlo method when a second order scheme is used instead of the Euler scheme which is, of course, of order one. Recall that a discretization scheme is said to be of second order when the quantity  $\sigma_\ell^2$ , introduced above, is of order  $m^{-2\ell}$ . For example, this is the case for the Milstein scheme ( see e.g. Kloeden and Platen [17] for more details on second order schemes). By the same reasoning as above, to achieve a given RMSE error for his multilevel estimator of order  $1/n^\alpha$ , Giles obtains an optimal choice of the parameters given by (see Theorem 3.1 of [8])

$$N_\ell = 2c_2 n^{2\alpha} \sqrt{T} \left( \frac{\sqrt{m} - 1}{\sqrt{m}} \right) \left( \frac{T}{m^\ell} \right)^{3/2} \quad \text{for } \ell \in \{0, \dots, L\} \quad \text{and} \quad L = \frac{\log n}{\log m}. \quad (2)$$

This choice leads to an optimal complexity for the multilevel Monte Carlo proportional to  $n^{2\alpha}$ .

In the present paper, we are interested in using Kebaier's approach [16] to get the optimal choice for the Multilevel Monte Carlo method. More precisely, our main result is a Lindeberg Feller central limit theorem for the Multilevel Monte Carlo Euler algorithm (see Theorem 5). In order to show this result, we first prove a stable law convergence theorem, for the Euler

scheme error on two consecutive levels  $m^{\ell-1}$  and  $m^\ell$ , of the type obtained in Jacod and Protter [15]. Indeed, we prove the following functional result (see Theorem 4)

$$\sqrt{\frac{m^\ell}{(m-1)T}}(X^{m^\ell} - X^{m^{\ell-1}}) \Rightarrow^{stably} U, \quad \text{as } \ell \rightarrow \infty,$$

where  $U$  is the same limit process given in Theorem 3.2 of Jacod and Protter [15]. In fact, their result, namely

$$\sqrt{\frac{m^\ell}{T}}(X^{m^\ell} - X) \Rightarrow^{stably} U, \quad \text{as } \ell \rightarrow \infty,$$

is not sufficient to prove our Theorem 5, since the multilevel Monte Carlo Euler method involves the error process  $X^{m^\ell} - X^{m^{\ell-1}}$  rather than  $X^{m^\ell} - X$ . Thanks to Theorem 5 we obtain a precise description for the choice of the parameters to run the multilevel Monte Carlo Euler method. Afterward, by a complexity analysis we obtain the optimal choice for the multilevel Monte Carlo Euler method. It turns out that for a total error of order  $1/n^\alpha$  the optimal parameters are given by

$$N_0 = n^{2\alpha}, \quad N_\ell = \frac{(m-1)Tn^{2\alpha} \log n}{m^\ell \log m}, \quad \text{for } \ell \in \{1, \dots, L\} \quad \text{and} \quad L = \frac{\log n}{\log m}. \quad (3)$$

This leads us to a complexity proportional to  $n^{2\alpha}(\log n)^2$  which is the same order obtained by Giles [8]. By comparing relations (1) and (3), we note that our optimal sequence of sample sizes  $(N_\ell)_{0 \leq \ell \leq L}$  does not depend on any given constant, since our approach is based on proving a central limit theorem and not on obtaining an upper bound for the variance of the algorithm. All these results are stated and proved in section 3.

In section 4, we investigate the application of this method to the pricing of Asian options. We proceed by approximating the integral in the payoff process using first the classical Riemann discretization scheme then the trapezoidal one. It was shown in Lapeyre and Temam [21] that these discretization schemes are both of second order and the associated weak error  $\varepsilon_n$  is of order  $n^{-1}$  (see section 4). At first, we prove two stable law convergence theorems, for the errors of both Riemann and trapezoidal schemes, on two consecutive levels  $m^{\ell-1}$  and  $m^\ell$  (see Theorem 6 and Theorem 7). We obtain a rate of convergence equal to  $m^\ell$  and the limit processes in both theorems are related to the one obtained by Kebaier [16] for the same setting. Then, we take advantage of this study to establish two new Lindeberg Feller central limit theorems (see Theorem 8 and Theorem 9). These results provide us a precise description for the choice of the parameters in the multilevel Monte Carlo method when used to price Asian options. In this context of second order schemes, the optimal sequence of sample sizes  $(N_\ell)_{0 \leq \ell \leq L}$  proposed by Giles [8] (see relation (2) with  $\alpha = 1$ ) does not satisfy the so called Lyapunov assumption of the Lindeberg Feller central limit theorem (see subsection 4.3). Indeed, Giles's analysis is only based on a control of the variance. However, our approach is based on proving a central limit theorem for the multilevel Monte Carlo method and we need in addition to satisfy a Lyapunov type condition that controls a moment of order greater than 2. Finally, we provide three possible choices of  $(N_\ell)_{0 \leq \ell \leq L}$  satisfying assumptions of the Lindeberg Feller central limit theorem and for which the optimal complexities can be closer to the order  $C_{MMC} = n^{2\alpha}$  but without reaching it (see subsection 4.4). Section 2 below is devoted to recall some useful stochastic limit theorems and to introduce our notations.

## 2 General framework

### 2.1 Preliminaries

We first recall basic facts about stable convergence. In the following we adopt the notation of Jacod and Protter [15]. Let  $X_n$  be a sequence of random variables with values in a Polish space  $E$ , all defined on the same probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  be an extension of  $(\Omega, \mathcal{F}, \mathbb{P})$ , and let  $X$  be an  $E$ -valued random variable on the extension. We say that  $(X_n)$  converges in law to  $X$  stably and write  $X_n \Rightarrow^{stably} X$ , if

$$\mathbb{E}(Uh(X_n)) \rightarrow \tilde{\mathbb{E}}(Uh(X))$$

for all  $h : E \rightarrow \mathbb{R}$  bounded continuous and all bounded random variable  $U$  on  $(\Omega, \mathcal{F})$ . This convergence, introduced by Rényi [23] and studied by Aldous and Eagelson [1], is obviously stronger than convergence in law that we will denote here by “ $\Rightarrow$ ”. According to section 2 of Jacod [14] and Lemma 2.1 of Jacod and Protter [15], we have the following results.

**Lemma 1** *let  $V_n$  and  $V$  be defined on  $(\Omega, \mathcal{F})$  with values in another metric space  $E'$ .*

$$\text{if } V_n \xrightarrow{\mathbb{P}} V, \quad X_n \Rightarrow^{stably} X \quad \text{then} \quad (V_n, X_n) \Rightarrow^{stably} (V, X)$$

*This result remains valid when  $V_n = V$  and conversely, if  $(V, X_n) \Rightarrow (V, X)$ , we can realize this limit as  $(V, X)$  with  $X$  defined on an extension of  $(\Omega, \mathcal{F}, \mathbb{P})$  and  $X_n \Rightarrow^{stably} X$  as soon as  $V$  generates the  $\sigma$ -field  $\mathcal{F}$ .*

Note that all this applies when  $X_n, X$  are  $\mathbb{R}^d$ -valued right-continuous and left-hand limited processes, where  $E = \mathcal{D}([0, T], \mathbb{R}^d)$  is equipped with the Skorokhod topology.

Now, we recall a result on the convergence of stochastic integrals formulated from Jacod and Protter [15]. This is a simplified version but it is sufficient for our study. Let  $X^n = (X^{n,i})_{1 \leq i \leq d}$  be a sequence of  $\mathbb{R}^d$ -valued continuous semimartingales with the decomposition

$$X_t^{n,i} = X_0^{n,i} + A_t^{n,i} + M_t^{n,i}, \quad 0 \leq t \leq T$$

where, for each  $n \in \mathbb{N}$  and  $1 \leq i \leq d$ ,  $A^{n,i}$  is a predictable process with finite variation, null at 0 and  $M^{n,i}$  is a martingale null at 0.

**Theorem 1** *Assume that the sequence  $(X^n)$  is such that*

$$\langle M^{n,i} \rangle_T + \int_0^T |dA_s^{n,i}|$$

*is tight. Let  $H^n$  and  $H$  be a sequence of adapted, right-continuous and left-hand limited processes all defined on the same filtered probability space. If  $(H^n, X^n) \Rightarrow (H, X)$  then  $X$  is a semimartingale with respect to the filtration generated by the limit process  $(H, X)$ , and we have  $(H^n, X^n, \int H^n dX^n) \Rightarrow (H, X, \int H dX)$ .*

## 2.2 The Euler scheme

Let  $X := (X_t)_{0 \leq t \leq T}$  be the process with values in  $\mathbb{R}^d$ , solution to

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d \quad (4)$$

where  $W = (W^1, \dots, W^q)$  is a  $q$ -dimensional Brownian motion on some given filtered probability space  $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ .  $(\mathcal{F}_t)_{t \geq 0}$  is the standard Brownian filtration.

The functions  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times q}$  are continuously differentiable and satisfy  $\exists C_T > 0; \forall x, y \in \mathbb{R}^d$  we have

$$\|b(x) - b(y)\| + \|\sigma(x) - \sigma(y)\| \leq C_T(\|y - x\|).$$

We consider the Euler continuous approximation  $X^n$  with step  $\delta = T/n$  given by:

$$dX_t^n = b(X_{\eta_n(t)})dt + \sigma(X_{\eta_n(t)})dW_t, \quad \eta_n(t) = [t/\delta]\delta.$$

It is well known that the Euler scheme satisfies the following properties (see for instance Faure [6] for more details)

$$\begin{aligned} \mathcal{P1}) \quad & \forall p > 1, \quad \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t - X_t^n|^p \right] \leq \frac{K_p(T)}{n^{p/2}}, \quad K_p(T) > 0. \\ \mathcal{P2}) \quad & \forall p > 1, \quad \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t|^p \right] + \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^n|^p \right] \leq K'_p(T), \quad K'_p(T) > 0. \end{aligned}$$

## 2.3 Stable convergence for the Euler scheme error

Now assume that

$$\varphi(X_t) = \begin{pmatrix} b_1(X_t) & \sigma_{11}(X_t) & \dots & \sigma_{1q}(X_t) \\ b_2(X_t) & \sigma_{21}(X_t) & \dots & \sigma_{2q}(X_t) \\ \vdots & \vdots & & \vdots \\ b_d(X_t) & \sigma_{d1}(X_t) & \dots & \sigma_{dq}(X_t) \end{pmatrix} \quad \text{and} \quad dY_t := \begin{pmatrix} dt \\ dW_t^1 \\ \vdots \\ dW_t^q \end{pmatrix}$$

then the S.D.E (4) becomes

$$dX_t = \varphi(X_t)dY_t = \sum_{j=0}^q \varphi_j(X_t)dY_t^j$$

where  $\varphi_j$  is  $j$ th column of the matrix  $\sigma$ , for  $1 \leq j \leq q$ , and  $\varphi_0 = b$ . The Euler continuous approximation  $X^n$  with step  $\delta = T/n$  is given by

$$dX_t^n = \varphi(X_{\eta_n(t)}^n)dY_t = \sum_{j=0}^q \varphi_j(X_{\eta_n(t)}^n)dY_t^j, \quad \eta_n(t) = [t/\delta]\delta. \quad (5)$$

The following result proven by Jacod and Protter [15] is an improvement of the result given by Kurtz and Protter [19].

**Theorem 2** *With the above notations we have*

$$\sqrt{\frac{n}{T}}(X^n - X) \Rightarrow^{stably} U$$

with  $(U_t)_{0 \leq t \leq T}$  the  $d$ -dimensional process satisfying

$$U_t = \frac{1}{\sqrt{2}} \sum_{i,j=1}^q Z_t \int_0^t Z_s^{-1} \dot{\varphi}_j(X_s) \varphi_i(X_s) dB_s^{ij}, \quad t \in [0, T], \quad (6)$$

where  $(Z_t)_{0 \leq t \leq T}$  is the  $\mathbb{R}^{d \times d}$  valued process solution of the linear equation

$$Z_t = I_d + \sum_{j=0}^q \int_0^t \dot{\varphi}_j(X_s) dY_s^j Z_s, \quad t \in [0, T],$$

$\dot{\varphi}_j$  is a  $d \times d$  matrix with  $(\dot{\varphi}_j)_{ik}$  is the partial derivative of  $\varphi_{ij}$  with respect to the  $k$ -th coordinate, and  $(B^{ij})_{1 \leq i,j \leq q}$  is a standard  $q^2$ -dimensional Brownian motion independent of  $W$ . This process is defined on an extension  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$  of the space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .

## 2.4 Central limit theorem for Monte Carlo Euler method

In many applications (in particular for the pricing of financial securities), the effective computation of  $\mathbb{E}f(X_T)$  is crucial. The Monte Carlo method consists of the following steps:

- Approximate the process  $(X_t)_{0 \leq t \leq T}$  by the Euler scheme  $(X_t^n)_{0 \leq t \leq T}$ , with step  $T/n$ , which can be simulated.
- Evaluate the expectation on the approximating process  $f(X_T^n)$  by the Monte Carlo method. In order to evaluate  $\mathbb{E}f(X_T^n)$  by the Monte Carlo method,  $N$  independent copies  $f(X_{T,i}^n)_{1 \leq i \leq N}$  of  $f(X_T^n)$  are sampled and the expectation is approximated by the following quantity

$$\hat{f}^{n,N} := \frac{1}{N} \sum_{i=1}^N f(X_{T,i}^n).$$

The approximation is affected by two types of errors. An analytical error given by

$$\varepsilon_n := \mathbb{E}f(X_T^n) - \mathbb{E}f(X_T)$$

and a statistical error  $\hat{f}^{n,N} - \mathbb{E}f(X_T^n)$ , controlled by the central limit Theorem and which is of order  $1/\sqrt{N}$ . An interesting problem (studied by Duffie and Glynn [5] and Kurtz and Protter [20]) is to find  $N$  as a function of  $n$  so that both errors are of the same order. Talay and Tubaro [24] prove that if  $f$  is sufficiently smooth, then  $\varepsilon_n \sim c/n$  with  $c$  a given constant. A similar result was proven by Kurtz and Protter [20] for a function  $f \in \mathcal{C}^3$ . The same result was extended by Bally and Talay [2] for a measurable function  $f$  but with a nondegeneracy condition of Hörmander type on the diffusion. In the context of possibly degenerate diffusions  $X$  and  $\mathcal{C}^1$ -functions  $f$ , Kebaier [16] prove that the rate of convergence of the discretization error  $\varepsilon_n$  can be  $1/n^\alpha$  for all values of  $\alpha \in [1/2, 1]$  (see Proposition 2.2 of [16]). The following result highlights the behavior of the global error in the classical Monte Carlo method. It can be proved in the same way as the limit theorem given in Duffie and Glynn [5].



**Theorem 3** *Let  $f$  be an  $\mathbb{R}^d$ -valued function satisfying*

$$(\mathcal{H}_f) \quad |f(x) - f(y)| \leq C(1 + |x|^p + |y|^p)|x - y|, \quad \text{for some } C, p > 0.$$

*Assume that  $\mathbb{P}(X_T \notin \mathcal{D}_f) = 0$ , where  $\mathcal{D}_f := \{x \in \mathbb{R}^d; f \text{ is differentiable at } x\}$ , and that for some  $\alpha \in [1/2, 1]$  we have*

$$(\mathcal{H}_{\varepsilon_n}) \quad \lim_{n \rightarrow \infty} n^\alpha \varepsilon_n = C_f(T, \alpha).$$

*Then*

$$n^\alpha \left( \frac{1}{n^{2\alpha}} \sum_{i=1}^{n^{2\alpha}} f(X_{T,i}^n) - \mathbb{E}f(X_T) \right) \Rightarrow \sigma \bar{G} + C_f(T, \alpha)$$

*with  $\sigma^2 = \text{Var}(f(X_T))$  and  $\bar{G}$  a standard normal.*

A functional version of this theorem, with  $\alpha = 1$  was proven by Kurtz and Protter [20] for a function  $f$  of class  $\mathcal{C}^3$ . One can interpret the theorem as follows. For a total error of order  $1/n^\alpha$  the minimal computation effort necessary to run the Monte Carlo algorithm is obtained for  $N = n^{2\alpha}$ . This leads to an optimal time complexity of the algorithm given by

$$C_{MC} = C \times (nN) = C \times n^{2\alpha+1}, \quad \text{with } C \text{ some positive constant.}$$

### 3 The Multilevel Monte Carlo Euler method

It is well known that the rate of convergence in the Monte Carlo method depends on the variance of  $f(X_T^n)$  where  $X_T^n$  is the Euler scheme of step  $T/n$ . This is a crucial point in the practical implementation. A large number of reduction of variance methods are used in practice. The multilevel algorithm proposes an iterative control variate reduction of variance that extends the the statistical Romberg method of Kebaier [16] (see also section 1 above). Its specificity is that the control variate is constructed in an iterative way by the Monte Carlo method using different time steps  $m^{-\ell}T$ ,  $\ell \in \{0, 1, \dots, L\}$  and  $m \in \mathbb{N} \setminus \{0, 1\}$  and such that  $m^L = n$ . Let us be more precise, it is clearly that

$$\mathbb{E}f(X_T^n) = \mathbb{E}f(X_T^1) + \sum_{\ell=1}^L \mathbb{E} \left( f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) \right). \quad (7)$$

The multilevel method is to estimate independently by the Monte Carlo method each of the expectations on the right-hand side. Hence, we approximate  $\mathbb{E}f(X_T^n)$  by

$$Q_n = \frac{1}{N_0} \sum_{k=1}^{N_0} f(X_{T,k}^1) + \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left( f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) \right). \quad (8)$$

The process  $(X_t^{m^\ell})_{0 \leq t \leq T}$  denotes the Euler scheme with time step  $m^{-\ell}T$  for  $\ell \in \{0, \dots, L\}$ , where  $L = \log n / \log m$ . Here, it is important to point out that all these  $L + 1$  Monte Carlo estimators have to be based on different, independent samples. However, for each  $k$  and  $\ell$  the simulations  $f(X_{T,k}^{m^\ell})$  and  $f(X_{T,k}^{m^{\ell-1}})$  come from the same Brownian path but with different time steps. The following result gives us a first description of the asymptotic behavior of the variance in the Multilevel Monte Carlo Euler method.

**Proposition 1** For a function  $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$  which is Lipschitz continuous of constant  $[f]_{lip}$  that is  $[f]_{lip} = \sup_{u \neq v} \frac{|f(u) - f(v)|}{\|u - v\|}$ , we have

$$Var(Q_n) = O\left(\sum_{\ell=0}^L N_\ell^{-1} m^{-\ell}\right). \quad (9)$$

**Proof :** We have

$$\begin{aligned} Var(Q_n) &= N_0^{-1} Var(f(X_T^1)) + \sum_{\ell=1}^L N_\ell^{-1} Var(f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})) \\ &\leq N_0^{-1} Var(f(X_T^1)) + 2 \sum_{\ell=1}^L N_\ell^{-1} \left( Var(f(X_T^{m^\ell}) - f(X_T)) + Var(f(X_T^{m^{\ell-1}}) - f(X_T)) \right) \\ &\leq N_0^{-1} Var(f(X_T^1)) + 2[f]_{lip} \sum_{\ell=1}^L N_\ell^{-1} \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^{m^\ell} - X_t|^2 + \sup_{0 \leq t \leq T} |X_t^{m^{\ell-1}} - X_t|^2 \right]. \end{aligned}$$

We complete the proof by using **P1**) on the strong convergence of the Euler scheme.  $\square$

The inequality (9) shows that the variance of  $Q_n$  depends on the choice of sample size  $N_\ell$ ,  $\ell \in \{0, 1, \dots, L\}$ . This variance can be smaller than the variance of  $f(X_T^n)$ , so that  $Q_n$  appears as a good candidate for the reduction of variance method.

The main result of this section is a Lindeberg Feller central limit theorem for the Multilevel Monte Carlo Euler algorithm (See Theorem 5 below). In order to prove this result, we need to prove first a stable law convergence theorem for the Euler scheme error. This is the aim of the following subsection.

### 3.1 Stable convergence

In what follows, we prove a stable law convergence theorem, for the Euler scheme error on two consecutive levels  $m^{\ell-1}$  and  $m^\ell$ , of the type obtained in Jacod and Protter [15] (see Theorem 2 above). Indeed, their result namely,

$$\sqrt{\frac{m^\ell}{T}}(X^{m^\ell} - X) \Rightarrow^{stably} U, \quad \text{as } \ell \rightarrow \infty,$$

is not sufficient to prove our Theorem 5 below, since the multilevel Monte Carlo Euler method involves the error process  $X^{m^\ell} - X^{m^{\ell-1}}$  rather than  $X^{m^\ell} - X$ . Note that the study of the error  $X^{m^\ell} - X^{m^{\ell-1}}$  as  $\ell \rightarrow \infty$  can be reduced to the study of the error  $X^{mn} - X^n$  as  $n \rightarrow \infty$ .

**Theorem 4** Under notations of Theorem 2, we have the following result

$$\sqrt{\frac{mn}{(m-1)T}}(X^{mn} - X^n) \Rightarrow^{stably} U,$$

where  $U$  is solution to equation (6) and  $m \in \mathbb{N} \setminus \{0, 1\}$ .

**Proof** : Consider the error process  $U^{mn,n} = (U_t^{mn,n})_{0 \leq t \leq T}$ , defined by

$$U_t^{mn,n} := X_t^{mn} - X_t^n, \quad t \in [0, T].$$

Combining relation (5), for both processes  $X^{mn}$  and  $X^n$ , together with a Taylor expansion yield us

$$dU_t^{mn,n} = \sum_{j=0}^q \dot{\varphi}_{t,j}^n (X_{\eta_{mn}(t)}^{mn} - X_{\eta_n(t)}^n) dY_t^j,$$

where

$$\dot{\varphi}_{t,j}^n = \int_0^1 \nabla \varphi_j (X_{\eta_n(t)}^n + \lambda(X_{\eta_{mn}(t)}^{mn} - X_{\eta_n(t)}^n)) d\lambda.$$

Therefore, the equation satisfied by  $U^n$  can be written as

$$U_t^{mn,n} = \int_0^t \sum_{j=0}^q \dot{\varphi}_{s,j}^n U_s^{mn,n} dY_s^j + G_t^{mn,n},$$

with

$$G_t^{mn,n} = \int_0^t \sum_{j=0}^q \dot{\varphi}_{s,j}^n (X_s^n - X_{\eta_n(s)}^n) dY_s^j - \int_0^t \sum_{j=0}^q \dot{\varphi}_{s,j}^n (X_s^{mn} - X_{\eta_{mn}(s)}^{mn}) dY_s^j.$$

In the following, let  $(Z_t^{mn,n})_{0 \leq t \leq T}$  be the  $\mathbb{R}^{d \times d}$  valued solution of

$$Z_t^{mn,n} = I_d + \int_0^t \left( \sum_{j=0}^q \dot{\varphi}_{s,j}^n dY_s^j \right) Z_s^{mn,n}.$$

Theorem 48 p.326 in [22], ensures the existence of the process  $((Z_t^{mn,n})^{-1})_{0 \leq t \leq T}$  solution to

$$(Z_t^{mn,n})^{-1} = I_d + \int_0^t (Z_s^{mn,n})^{-1} \sum_{j=1}^q (\dot{\varphi}_{s,j}^n)^2 ds - \int_0^t (Z_s^{mn,n})^{-1} \sum_{j=0}^q \dot{\varphi}_{s,j}^n dY_s^j.$$

Thanks to theorem 56 p. 33 in the same reference [22], we get

$$\begin{aligned} U_t^{mn,n} = Z_t^{mn,n} \Big\{ & \int_0^t (Z_s^{mn,n})^{-1} dG_s^{mn,n} - \int_0^t (Z_s^{mn,n})^{-1} \sum_{j=1}^q (\dot{\varphi}_{s,j}^n)^2 (X_s^n - X_{\eta_n(s)}^n) ds \\ & + \int_0^t (Z_s^{mn,n})^{-1} \sum_{j=1}^q (\dot{\varphi}_{s,j}^n)^2 (X_s^{mn} - X_{\eta_{mn}(s)}^{mn}) ds \Big\}. \end{aligned}$$

Since the increments of the Euler scheme satisfy

$$X_s^n - X_{\eta_n(s)}^n = \sum_{i=0}^q \bar{\varphi}_{s,i}^n (Y_s^i - Y_{\eta_n(s)}^i) \quad \text{and} \quad X_s^{mn} - X_{\eta_{mn}(s)}^{mn} = \sum_{i=0}^q \bar{\varphi}_{s,i}^{mn} (Y_s^i - Y_{\eta_{mn}(s)}^i),$$

with  $\bar{\varphi}_{s,i}^n = \varphi_i(X_{\eta_n(s)}^n)$  and  $\bar{\varphi}_{s,i}^{mn} = \varphi_i(X_{\eta_{mn}(s)}^{mn})$ , it is easy to check that

$$\begin{aligned} U_t^{mn,n} = & \sum_{i,j=1}^q Z_t^{mn,n} \int_0^t H_s^{i,j,mn,n} (Y_s^i - Y_{\eta_n(s)}^i) dY_s^j + R_{t,1}^{mn,n} + R_{t,2}^{mn,n} \\ & - \sum_{i,j=1}^q Z_t^{mn,n} \int_0^t \tilde{H}_s^{i,j,mn,n} (Y_s^i - Y_{\eta_{mn}(s)}^i) dY_s^j - \tilde{R}_{t,1}^{mn,n} - \tilde{R}_{t,2}^{mn,n} \quad (10) \end{aligned}$$

with

$$R_{t,1}^{mn,n} = \sum_{i=0}^q Z_t^{mn,n} \int_0^t K_s^{i,mn,n} (Y_s^i - Y_{\eta_n(s)}^i) ds, \quad R_{t,2}^{mn,n} = \sum_{j=1}^q Z_t^{mn,n} \int_0^t H_s^{0,j,mn,n} (s - \eta_n(s)) dY_s^j,$$

and

$$\tilde{R}_{t,1}^{mn,n} = \sum_{i=0}^q Z_t^{mn,n} \int_0^t \tilde{K}_s^{i,mn,n} (Y_s^i - Y_{\eta_{mn}(s)}^i) ds, \quad \tilde{R}_{t,2}^{mn,n} = \sum_{j=1}^q Z_t^{mn,n} \int_0^t \tilde{H}_s^{0,j,mn,n} (s - \eta_{mn}(s)) dY_s^j.$$

where, for  $(i, j) \in \{0, \dots, q\} \times \{1, \dots, q\}$ ,

$$K_s^{i,mn,n} = (Z_s^{mn,n})^{-1} \left( \dot{\varphi}_{s,0}^n \bar{\varphi}_{s,i}^n - \sum_{j=1}^q (\dot{\varphi}_{s,j}^n)^2 \bar{\varphi}_{s,i}^n \right), \quad H_s^{i,j,mn,n} = (Z_s^{mn,n})^{-1} \dot{\varphi}_{s,j}^n \bar{\varphi}_{s,i}^n,$$

and

$$\tilde{K}_s^{i,mn,n} = (Z_s^{mn,n})^{-1} \left( \dot{\varphi}_{s,0}^n \bar{\varphi}_{s,i}^{mn} - \sum_{j=1}^q (\dot{\varphi}_{s,j}^n)^2 \bar{\varphi}_{s,i}^{mn} \right), \quad \tilde{H}_s^{i,j,mn,n} = (Z_s^{mn,n})^{-1} \dot{\varphi}_{s,j}^n \bar{\varphi}_{s,i}^{mn}.$$

Now, let us introduce  $Z_t = D_x X_t$  solution to

$$Z_t = I_d + \int_0^t \sum_{j=0}^q (\dot{\varphi}_{s,j} dY_s^j) Z_s, \quad \text{with } \dot{\varphi}_{t,j} = \nabla \varphi_j(X_t).$$

Moreover,  $((Z_t)^{-1})_{0 \leq t \leq T}$  exists and satisfies the following explicit linear stochastic differential equation

$$(Z_t)^{-1} = I_d + \int_0^t (Z_s)^{-1} \sum_{j=1}^q (\dot{\varphi}_{s,j})^2 ds - \int_0^t (Z_s)^{-1} \sum_{j=0}^q \dot{\varphi}_{s,j} dY_s^j.$$

Note that using the same techniques as in the proof of existence and uniqueness for stochastic differential equations with Lipschitz coefficients (i.e. Gronwall inequality), we obtain that for any  $p \geq 1$  and for any  $t \in [0, T]$ ,  $Z_t^{mn,n}$ ,  $Z_t$ ,  $(Z_t^{mn,n})^{-1}$ ,  $(Z_t)^{-1} \in L^p$  and

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \sup_{0 \leq t \leq T} |Z_t^{mn,n} - Z_t|^p \right] = 0, \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[ \sup_{0 \leq t \leq T} |(Z_t^{mn,n})^{-1} - (Z_t)^{-1}|^p \right] = 0. \quad (11)$$

Furthermore, in relation (10), one can replace respectively  $H_s^{i,j,mn,n}$  and  $\tilde{H}_s^{i,j,mn,n}$  by their common limit

$$H_s^{i,j} = (Z_s)^{-1} \dot{\varphi}_{s,j} \bar{\varphi}_{s,i}, \text{ with } \dot{\varphi}_{s,j} = \nabla \varphi_j(X_s) \text{ and } \bar{\varphi}_{s,i} = \varphi_i(X_s).$$

So that, relation (10) becomes

$$U_t^{mn,n} = \sum_{i,j=1}^q Z_t^{mn,n} \int_0^t H_s^{i,j} (Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i) dY_s^j + R_t^{mn,n}, \quad (12)$$

with

$$R_t^{mn,n} = R_{t,1}^{mn,n} + R_{t,2}^{mn,n} + R_{t,3}^{mn,n} - \tilde{R}_{t,1}^{mn,n} - \tilde{R}_{t,2}^{mn,n} - \tilde{R}_{t,3}^{mn,n}$$

where  $R_{t,i}^{mn,n}$  and  $\tilde{R}_{t,i}^{mn,n}$ ,  $i \in \{1, 2\}$ , are introduced by relation (10) and

$$\begin{aligned} R_{t,3}^{mn,n} &= \sum_{i,j=1}^q Z_t^{mn,n} \int_0^t (H_s^{i,j,mn,n} - H_s^{i,j}) (Y_s^i - Y_{\eta_n(s)}^i) dY_s^j \\ \tilde{R}_{t,3}^{mn,n} &= \sum_{i,j=1}^q Z_t^{mn,n} \int_0^t (\tilde{H}_s^{i,j,mn,n} - H_s^{i,j}) (Y_s^i - Y_{\eta_{mn}(s)}^i) dY_s^j, \end{aligned}$$

The remainder term process  $R^{mn,n}$  vanishes with rate  $\sqrt{n}$  in probability. More precisely, we have the following convergence result.

**Lemma 2** *The rest term introduced in relation (12) satisfies  $\sup_{0 \leq t \leq T} |\sqrt{n} R_t^{mn,n}|$  converges to zero in probability as  $n$  tends to infinity.*

For the reader convenience, the proof of this lemma is postponed to the end of the current subsection.

The task is now to study the asymptotic behavior of the process given by relation (12)

$$\sum_{i,j=1}^q \sqrt{n} Z_t^{mn,n} \int_0^t H_s^{i,j} (Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i) dY_s^j.$$

In order to study this process, we introduce the martingale process,

$$M_t^{n,i,j} = \int_0^t (Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i) dY_s^j, \quad (i, j) \in \{1, \dots, q\}^2,$$

and we proceed to a preliminary calculus of the expectation of its bracket. Let  $(i, j)$  and  $(i', j') \in \{1, \dots, q\}^2$ , we have

- for  $j \neq j'$ , the bracket  $\langle M^{n,i,j}, M^{n,i',j'} \rangle = 0$
- for  $j = j'$  and  $i \neq i'$ ,  $\mathbb{E} \langle M^{n,i,j}, M^{n,i',j} \rangle = 0$

- for  $j = j'$  and  $i = i'$ ,  $\mathbb{E}\langle M^{n,i,j} \rangle_t = \int_0^t (\eta_{mn}(s) - \eta_n(s)) ds$ ,  $t \in [0, T]$  and we have

$$\begin{aligned}
\mathbb{E}(\langle M^{n,i,j} \rangle_t) &= \int_0^{\eta_n(t)} (\eta_{mn}(s) - \eta_n(s)) ds + O\left(\frac{1}{n^2}\right) \\
&= \sum_{\ell=0}^{m-1} \sum_{k=0}^{[t/\delta]-1} \int_{(mk+\ell)\delta/m}^{(mk+\ell+1)\delta/m} (\eta_{mn}(s) - \eta_n(s)) ds + O\left(\frac{1}{n^2}\right) \\
&= \sum_{\ell=0}^{m-1} \sum_{k=0}^{[t/\delta]-1} \frac{\delta^2}{m} \left( \frac{mk+\ell}{m} - k \right) + O\left(\frac{1}{n^2}\right) = \frac{(m-1)\delta^2}{2m} [t/\delta] + O\left(\frac{1}{n^2}\right) \\
&= \frac{(m-1)T}{2mn} t + O\left(\frac{1}{n^2}\right). \tag{13}
\end{aligned}$$

Having disposed of this preliminary evaluations, we can now study the stable convergence of  $\left(\sqrt{\frac{2mn}{(m-1)T}} M^{n,i,j}\right)_{1 \leq i,j \leq q}$ . By virtue of Theorem 2-1 of [14], we need to study the asymptotic behavior of both brackets  $n\langle M^{n,i,j}, M^{n,i',j'} \rangle_t$  and  $\sqrt{n}\langle M^{n,i,j}, Y^{j'} \rangle_t$ , for all  $t \in [0, T]$  and all  $(i, j, i', j') \in \{1, \dots, q\}^4$ . The case  $j \neq j'$  is obvious and we only proceed to prove that

- for  $j = j'$ ,  $\sqrt{n}\langle M^{n,i,j}, Y^j \rangle_t \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0$ , for all  $t \in [0, T]$ .
- for  $j = j'$  and  $i \neq i'$ ,  $n\langle M^{n,i,j}, M^{n,i',j} \rangle_t \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0$ , for all  $t \in [0, T]$ .
- for  $j = j'$  and  $i = i'$ ,  $n\langle M^{n,i,j} \rangle_t \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \frac{(m-1)T}{2m} t$ , for all  $t \in [0, T]$ .

For the first point, we consider the  $L^2$  convergence

$$\begin{aligned}
\mathbb{E}\langle M^{n,i,j}, Y^j \rangle_t^2 &= \mathbb{E} \left( \int_0^t (Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i) ds \right)^2 \\
&= \int_0^t \int_0^t \mathbb{E}((Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i)(Y_{\eta_{mn}(u)}^i - Y_{\eta_n(u)}^i)) ds du \\
&= 2 \int_{0 < s < u < t} g(s, u) ds du
\end{aligned}$$

with

$$g(s, u) = \eta_{mn}(s) \wedge \eta_{mn}(u) - \eta_{mn}(s) \wedge \eta_n(u) - \eta_n(s) \wedge \eta_{mn}(u) + \eta_n(s) \wedge \eta_n(u). \tag{14}$$

It is worthy to note that

$$\eta_n(s) \leq \eta_{mn}(s) \leq s \leq \eta_n(u) \leq \eta_{mn}(u) \leq u, \quad \forall s \leq \eta_n(u). \tag{15}$$

Hence  $g(s, u) = 0$ , for  $s \leq \eta_n(u)$ ,  $g(s, u) = \eta_{mn}(s) - \eta_n(s)$ , for  $\eta_n(u) < s < u$ , and

$$\mathbb{E} \langle M^{n,i,j}, Y^j \rangle_t^2 = 2 \int_{0 < \eta_n(u) < s < u < t} (\eta_{mn}(s) - \eta_n(s)) ds du \leq 2 \frac{T}{n} \int_0^t (u - \eta_n(u)) du \leq 2 \frac{T^2}{n^2} t.$$

This yields the desired result. Concerning the second point, the  $L^2$  norm is given by

$$\begin{aligned}\mathbb{E}\langle M^{n,i,j}, M^{n,i',j} \rangle_t^2 &= \mathbb{E} \left( \int_0^t (Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i)(Y_{\eta_{mn}(s)}^{i'} - Y_{\eta_n(s)}^{i'}) ds \right)^2 \\ &= \int_0^t \int_0^t (\mathbb{E}((Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i)(Y_{\eta_{mn}(u)}^i - Y_{\eta_n(u)}^i)))^2 ds du \\ &= 2 \int_{0 < s < u < t} g(s, u)^2 ds du,\end{aligned}$$

with the same function  $g$  given in relation (14). By properties of  $g$  developed above, we have in the same manner

$$\mathbb{E}\langle M^{n,i,j}, M^{n,i',j} \rangle_t^2 = 2 \int_{0 < \eta_n(u) < s < u < t} (\eta_{mn}(s) - \eta_n(s))^2 ds du \leq 2 \frac{T^3}{n^3} t,$$

which proves our claim. For the last point, that is the essential one, taking into account the development of  $\mathbb{E}\langle M^{n,i,j} \rangle_t$  given by relation (13) we obtain

$$\mathbb{E} \left( n \langle M^{n,i,j} \rangle_t - \frac{(m-1)T}{2m} t \right)^2 = n^2 \mathbb{E}\langle M^{n,i,j} \rangle_t^2 - \frac{(m-1)^2 T^2}{4m^2} t^2 + O\left(\frac{1}{n}\right). \quad (16)$$

Otherwise, we have

$$\begin{aligned}\mathbb{E}\langle M^{n,i,j} \rangle_t^2 &= \mathbb{E} \left( \int_0^t (Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i)^2 ds \right)^2 \\ &= \int_0^t \int_0^t \mathbb{E}((Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i)^2 (Y_{\eta_{mn}(u)}^i - Y_{\eta_n(u)}^i)^2) ds du \\ &= 2 \int_{0 < s < u < t} h(s, u) ds du\end{aligned} \quad (17)$$

with

$$h(s, u) = \mathbb{E}((Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i)^2 (Y_{\eta_{mn}(u)}^i - Y_{\eta_n(u)}^i)^2). \quad (18)$$

On the one hand, for  $s \leq \eta_n(u)$ , by property (15) and since the increments  $Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i$  and  $Y_{\eta_{mn}(u)}^i - Y_{\eta_n(u)}^i$  are independent, it follows immediately that

$$h(s, u) = (\eta_{mn}(s) - \eta_n(s))(\eta_{mn}(u) - \eta_n(u)).$$

On the other hand, in relation (18) we use the Cauchy-Schwartz inequality to get  $h(s, u) = O(\frac{1}{n^2})$  and this yields

$$\int_{0 < \eta_n(u) < s < u < t} h(s, u) ds du = O\left(\frac{1}{n^3}\right).$$

Now, noting that  $(\eta_{mn}(s) - \eta_n(s))(\eta_{mn}(u) - \eta_n(u)) = O(\frac{1}{n^2})$ , relation (17) becomes

$$\begin{aligned}\mathbb{E}(\langle M^{n,i,j} \rangle_t^2) &= 2 \int_{0 < s < u < t} (\eta_{mn}(s) - \eta_n(s))(\eta_{mn}(u) - \eta_n(u)) ds du + O\left(\frac{1}{n^3}\right) \\ &= \left( \int_0^t (\eta_{mn}(s) - \eta_n(s)) ds \right)^2 + O\left(\frac{1}{n^3}\right).\end{aligned}$$

Once again thanks to the development of  $\mathbb{E}(\langle M^{n,i,j} \rangle_t)$  given by relation (13), we deduce that

$$\mathbb{E} \langle M^{n,i,j} \rangle_t^2 = \frac{(m-1)^2 T^2}{4m^2 n^2} t^2 + O\left(\frac{1}{n^3}\right). \quad (19)$$

Combining relations (16) and (19), we deduce the convergence in  $L^2$  of  $n \langle M^{n,i,j} \rangle_t$  towards  $\frac{(m-1)T}{2m}t$ . Hence  $\left(\sqrt{\frac{2mn}{(m-1)T}} M^{n,i,j}\right)_{1 \leq i,j \leq q}$  converges in law stably to a standard  $q^2$ -dimensional Brownian motion  $(B^{ij})_{1 \leq i,j \leq q}$  independent of  $W$ . Consequently, by Lemma 1 and Theorem 1, we obtain

$$\left(\sqrt{\frac{mn}{(m-1)T}} \int_0^t H_s^{i,j} (Y_{\eta_{mn}(s)}^i - Y_{\eta_n(s)}^i) dY_s^j, t \geq 0\right)_{1 \leq i,j \leq q} \Rightarrow^{stably} \left(\int_0^t H_s^{i,j} \frac{dB_s^{ij}}{\sqrt{2}}, t \geq 0\right)_{1 \leq i,j \leq q}$$

Finally, we complete the proof using relations (11), (12), Lemma 2 and once again Lemma 1 to obtain

$$\sqrt{\frac{mn}{(m-1)T}} U^{mn,n} \Rightarrow^{stably} U, \quad \text{where} \quad U_t = \frac{1}{\sqrt{2}} \sum_{i,j=1}^q Z_t \int_0^t H_s^{i,j} dB_s^{ij}.$$

□

**Proof of lemma 2 :** At first, we prove the uniform probability convergence toward zero of the normalized rest terms  $\sqrt{n} R_{t,i}^{mn,n}$  for  $i \in \{1, 2\}$ . The convergence of  $\sqrt{n} \tilde{R}_{t,i}^{mn,n}$   $i \in \{1, 2\}$  is a straightforward consequence of the previous one. The main part of these rest terms can be represented as integrals with respect to three types of supermartingales that can be classified through the following three cases

$$D_t^{n,0,0} = \sqrt{n} \int_0^t (s - \eta_n(s)) ds, \quad D_t^{n,i,0} = \sqrt{n} \int_0^t (Y_s^i - Y_{\eta_n(s)}^i) ds, \quad M_t^{n,0,j} = \sqrt{n} \int_0^t (s - \eta_n(s)) dY_s^j,$$

where  $(i, j) \in \{1, \dots, q\}^2$  and  $t \in [0, T]$ . In the first case the supermartingale is deterministic of finite variation and its total variation on the interval  $[0, T]$  has the following expression

$$\int_0^T |dD_t^{n,0,0}| = \sqrt{n} \int_0^T (s - \eta_n(s)) ds \leq \frac{T^2}{\sqrt{n}}.$$

So, the process  $D^{n,0,0}$  converges to 0 and is tight. In the second case, for  $i \in \{1, \dots, q\}$ , the supermartingale is also of finite variation and its total variation on the interval  $[0, T]$  has the following expression

$$\int_0^T |dD_t^{n,i,0}| = \sqrt{n} \int_0^T |Y_s^i - Y_{\eta_n(s)}^i| ds.$$

It is clear that  $\sup_n \mathbb{E} \left( \int_0^T |dD_s^{n,i,0}| \right) < \infty$  which ensures the tightness of the process  $D^{n,i,0}$ . Therefore, we only need to establish the convergence of  $D_t^{n,i,0}$  towards 0 in  $L^2(\Omega)$ , for  $t \in [0, T]$ . In fact, we have

$$\mathbb{E} ((D_t^{n,i,0})^2) = 2n \int_{0 < s < u < t} \mathbb{E} ((Y_s^i - Y_{\eta_n(s)}^i)(Y_u^i - Y_{\eta_n(u)}^i)) ds du.$$



When  $s \leq \eta_n(u)$ , we have  $\eta_n(s) \leq s \leq \eta_n(u) \leq u$  and by independence of the Brownian motion increments, we deduce that the integrand term is equal to 0. Otherwise, when  $s \geq \eta_n(u)$ , we apply the Cauchy Schwartz inequality to get

$$\mathbb{E}((D_t^{n,i,0})^2) \leq 2T \int_0^t (u - \eta_n(u)) du \leq 2 \frac{T^2}{n} t.$$

It follows from all these that  $D^{n,i,0} \Rightarrow 0$ . In the last case, for  $j \in \{1, \dots, q\}$ , the process  $M_t^{n,0,j}$  is a square integrable martingale and its bracket has the following expression

$$\langle M^{n,0,j} \rangle_T = n \int_0^T (s - \eta_n(s))^2 ds \leq \frac{T^3}{n}.$$

It is clear that  $\sup_n \mathbb{E} \langle M^{n,0,j} \rangle_T < \infty$ , so we deduce the tightness of the process  $\langle M^{n,0,j} \rangle$  and the convergence  $M^{n,0,j} \Rightarrow 0$ .

Now thanks to property  $\mathcal{P}1$ ) and relation (11), it is easy to check that the integrand processes  $K_s^{i,mn,n}$  and  $H_s^{0,j,mn,n}$ , introduced in relation (10), converge uniformly in probability to their respective limits  $K_s^i = (Z_s)^{-1} \left( \dot{\varphi}_{s,0} \bar{\varphi}_{s,i} - \sum_{j=1}^q (\dot{\varphi}_{s,j})^2 \bar{\varphi}_{s,i} \right)$  and  $H_s^{0,j} = (Z_s)^{-1} \dot{\varphi}_{s,j} \bar{\varphi}_{s,i}$ , where  $\dot{\varphi}_{s,j} = \nabla \varphi_j(X_s)$  and  $\bar{\varphi}_{s,i} = \varphi_i(X_s)$ . Therefore, by Theorem 1 we deduce that the integral processes given by

$$\sqrt{n} \int_0^t K_s^{i,mn,n} (Y_s^i - Y_{\eta_n(s)}^i) ds \quad \text{and} \quad \sqrt{n} \int_0^t H_s^{0,j,mn,n} (s - \eta_n(s)) dY_s^j$$

vanish. Consequently, we conclude using relation (11) that  $\sqrt{n} R_i^{mn,n} \Rightarrow 0$  for  $i \in \{1, 2\}$ .

We now proceed to prove that  $R_3^{mn,n} \Rightarrow 0$ . The convergence of the process  $\tilde{R}_3^{mn,n}$  toward 0 is obviously obtained from the previous one. The main part of this rest term can be represented as a stochastic integral with respect to the martingale process given by

$$N_t^{n,i,j} = \sqrt{n} \int_0^t (Y_s^i - Y_{\eta_n(s)}^i) dY_s^j,$$

with  $(i, j) \in \{1, \dots, q\} \times \{1, \dots, q\}$ . It was proven in Jacod and Protter [15] that

$$\sqrt{\frac{n}{T}} N^{n,i,j} \Rightarrow^{stably} \frac{B^{ij}}{\sqrt{2}},$$

where  $(B^{ij})_{1 \leq i,j \leq q}$  is a standard  $q^2$ -dimensional Brownian motion defined on an extension  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \mathbb{P})$  of the space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , which is independent of  $W$ . Thanks to property  $\mathcal{P}1$ ) and relation (11), the integrand process  $H^{i,j,mn,n} - H^{i,j} \Rightarrow 0$  and once again by Theorem 1 we deduce that the integral processes given by

$$\sqrt{n} \int_0^t (H_s^{i,j,mn,n} - H_s^{i,j}) (Y_s^i - Y_{\eta_n(s)}^i) dY_s^j$$

vanish. All this allows us to conclude using relation (11).  $\square$

### 3.2 Central limit theorem

Let us recall that the multilevel Monte Carlo method uses information from a sequence of computations with decreasing step sizes and approximates the quantity  $\mathbb{E}f(X_T)$  by

$$Q_n = \frac{1}{N_0} \sum_{k=1}^{N_0} f(X_{T,k}^1) + \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left( f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) \right), \quad m \in \mathbb{N} \setminus \{0, 1\} \quad \text{and} \quad L = \frac{\log n}{\log m}.$$

In the same way as in the case of a crude Monte Carlo estimation, let us assume that the discretization error

$$\varepsilon_n = \mathbb{E}f(X_T^n) - \mathbb{E}f(X_T)$$

is of order  $1/n^\alpha$  for any  $\alpha \in [1/2, 1]$ . Taking advantage from the limit theorem proven in the above section, we are now able to establish a central limit theorem of Lindeberg Feller type on the multilevel Monte Carlo Euler method. To do so, we introduce a real sequence  $(a_\ell)_{\ell \geq 1}$  of positive terms such that

$$(\mathcal{W}) \quad \lim_{L \rightarrow \infty} \sum_{\ell=1}^L a_\ell = \infty \quad \text{and} \quad \lim_{L \rightarrow \infty} \frac{1}{\left( \sum_{\ell=1}^L a_\ell \right)^{p/2}} \sum_{\ell=1}^L a_\ell^{p/2} = 0, \quad \text{for } p > 2.$$

and we assume that the sample size  $N_\ell$  depends on the rest of parameters by the relation

$$N_0 = n^{2\alpha}, \quad N_\ell = \frac{n^{2\alpha}(m-1)T}{m^\ell a_\ell} \sum_{\ell=1}^L a_\ell, \quad \ell \in \{1, \dots, L\} \quad \text{and} \quad L = \frac{\log n}{\log m}. \quad (20)$$

We choose this form for  $N_\ell$  because:

- it is a generic form covering our different studies for both first and second order discretization schemes (see subsection 4.3 below),
- it allows us a straightforward use of Toeplitz lemma that is a crucial tool used in the proof of our central limit theorem.

We can now state the analogue of Theorem 3 in our setting.

**Theorem 5** *Let  $f$  be an  $\mathbb{R}^d$ -valued function satisfying*

$$(\mathcal{H}_f) \quad |f(x) - f(y)| \leq C(1 + |x|^p + |y|^p)|x - y|, \quad \text{for some } C, p > 0.$$

*Assume that  $\mathbb{P}(X_T \notin \mathcal{D}_f) = 0$ , where  $\mathcal{D}_f := \{x \in \mathbb{R}^d; f \text{ is differentiable at } x\}$ , and that for some  $\alpha \in [1/2, 1]$  we have*

$$(\mathcal{H}_{\varepsilon_n}) \quad \lim_{n \rightarrow \infty} n^\alpha \varepsilon_n = C_f(T, \alpha).$$

*Then, for the choice of  $N_\ell$ ,  $\ell \in \{0, 1, \dots, L\}$  given by equation (20), we have*

$$n^\alpha (Q_n - \mathbb{E}(f(X_T))) \Rightarrow \mathcal{N}(C_f(T, \alpha), \sigma^2)$$

*with  $\sigma^2 = \text{Var}(f(X_T^1)) + \tilde{\text{Var}}(\nabla f(X_T)U_T)$  and  $\mathcal{N}(C_f(T, \alpha), \sigma^2)$  denotes a normal distribution.*

**Proof :** To simplify our notations we give the proof for  $\alpha = 1$ , the case  $\alpha \in [1/2, 1)$  is a straightforward deduction. Combining relations (7) and (8) together we get

$$Q_n - \mathbb{E}(f(X_T)) = \hat{Q}_n^1 + \hat{Q}_n^2 + \varepsilon_n,$$

where

$$\begin{aligned}\hat{Q}_n^1 &= \frac{1}{N_0} \sum_{k=1}^{N_0} (f(X_{T,k}^1) - \mathbb{E}(f(X_T^1))) \\ \hat{Q}_n^2 &= \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left( f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) - \mathbb{E} \left( f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) \right) \right).\end{aligned}$$

Using the assumption  $(\mathcal{H}_{\varepsilon_n})$  we obviously obtain the term  $C_f(T, \alpha)$  in the limit. Taking  $N_0 = n^2$ , we can apply the classical central limit theorem to  $n\hat{Q}_n^1$ . Then we have

$$n\hat{Q}_n^1 \Rightarrow \mathcal{N}(0, \text{Var}(f(X_T^1))).$$

Finally, we have only to study the convergence of  $n\hat{Q}_n^2$  and we will conclude by establishing

$$n\hat{Q}_n^2 \Rightarrow \mathcal{N}(0, \tilde{\text{Var}}(\nabla f(X_T).U_T)).$$

To do so, we plan to use the Lindeberg Feller theorem [10] with Lyapunov condition. We set

$$X_{n,\ell} := \frac{n}{N_\ell} \sum_{k=1}^{N_\ell} Z_{T,k}^{m^\ell, m^{\ell-1}} \text{ and } Z_{T,k}^{m^\ell, m^{\ell-1}} := f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) - \mathbb{E} \left( f(X_{T,K}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}}) \right).$$

In other words, we will check the following conditions :

- $\lim_{n \rightarrow \infty} \sum_{\ell=1}^L \mathbb{E}(X_{n,\ell})^2 = \tilde{\text{Var}}(\nabla f(X_T).U_T)$
- (Lyapunov condition) there exists  $p > 2$  such that  $\lim_{n \rightarrow \infty} \sum_{\ell=1}^L \mathbb{E}|X_{n,\ell}|^p = 0$ .

For the first one, we have

$$\begin{aligned}\sum_{\ell=1}^L \mathbb{E}(X_{n,\ell})^2 &= \sum_{\ell=1}^L \text{Var}(X_{n,\ell}) = \sum_{\ell=1}^L \frac{n^2}{N_\ell} \text{Var} \left( Z_{T,1}^{m^\ell, m^{\ell-1}} \right) \\ &= \frac{1}{\sum_{\ell=1}^L a_\ell} \sum_{\ell=1}^L a_\ell \frac{m^\ell}{(m-1)T} \text{Var} \left( Z_{T,1}^{m^\ell, m^{\ell-1}} \right).\end{aligned}\quad (21)$$

Otherwise, since  $\mathbb{P}(X_T \notin \mathcal{D}_f) = 0$ , applying the Taylor expansion theorem twice we get

$$\begin{aligned}f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) &= \nabla f(X_T).U_T^{m^\ell, m^{\ell-1}} + \\ &\quad (X_T^{m^\ell} - X_T) \varepsilon(X_T, X_T^{m^\ell} - X_T) - (X_T^{m^{\ell-1}} - X_T) \varepsilon(X_T, X_T^{m^{\ell-1}} - X_T).\end{aligned}$$

with  $\varepsilon(X_T, X_T^{m^\ell} - X_T) \xrightarrow{\mathbb{P}} 0$  and  $\varepsilon(X_T, X_T^{m^{\ell-1}} - X_T) \xrightarrow{\mathbb{P}} 0$ . By Theorem 2 we get the tightness of  $\sqrt{\frac{m^\ell}{(m-1)T}}(X_T^{m^\ell} - X_T)$  and  $\sqrt{\frac{m^\ell}{(m-1)T}}(X_T^{m^{\ell-1}} - X_T)$  and we deduce

$$\sqrt{\frac{m^\ell}{(m-1)T}} \left( (X_T^{m^\ell} - X_T) \varepsilon(X_T, X_T^{m^\ell} - X_T) - (X_T^{m^{\ell-1}} - X_T) \varepsilon(X_T, X_T^{m^{\ell-1}} - X_T) \right) \xrightarrow{\mathbb{P}} 0.$$

So, according to lemma 1 and Theorem 4 and since  $\nabla f(X_T^{m^{\ell-1}}) \xrightarrow{\mathbb{P}} \nabla f(X_T)$  we conclude that

$$\sqrt{\frac{m^\ell}{(m-1)T}} \left( f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) \right) \Rightarrow^{stably} \nabla f(X_T).U_T, \text{ as } \ell \rightarrow \infty. \quad (22)$$

Using  $(\mathcal{H}_f)$  it follows from property  $\mathcal{P}1)$  that

$$\forall \varepsilon > 0, \quad \sup_{\ell} \mathbb{E} \left| \sqrt{\frac{m^\ell}{(m-1)T}} \left( f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) \right) \right|^{2+\varepsilon} < \infty.$$

We deduce using relation (22) that

$$\mathbb{E} \left( \sqrt{\frac{m^\ell}{(m-1)T}} \left( f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) \right) \right)^k \rightarrow \tilde{\mathbb{E}} \left( \nabla f(X_T).U_T \right)^k < \infty \text{ for } k \in \{1, 2\}.$$

Consequently,

$$\frac{m^\ell}{(m-1)T} \text{Var}(Z_{T,1}^{m^\ell, m^{\ell-1}}) \longrightarrow \tilde{\text{Var}}(\nabla f(X_T).U_T) < \infty.$$

Hence combining this result with relation (21), we obtain the first condition using Toeplitz lemma. Concerning the second one, by Burkholder's inequality and elementary computations, we get for  $p > 2$

$$\mathbb{E}|X_{n,\ell}|^p = \frac{n^p}{N_\ell^p} \mathbb{E} \left| \sum_{\ell=1}^{N_\ell} Z_{T,1}^{m^\ell, m^{\ell-1}} \right|^p \leq C_p \frac{n^p}{N_\ell^{p/2}} \mathbb{E} \left| Z_{T,1}^{m^\ell, m^{\ell-1}} \right|^p,$$

where  $C_p$  is a numerical constant depending on  $p$  only. Otherwise, property  $\mathcal{P}1)$  ensures the existence of a constant  $K_p > 0$  such that

$$\left| \mathbb{E} Z_{T,1}^{m^\ell, m^{\ell-1}} \right|^p \leq \frac{K_p}{m^{p\ell/2}}.$$

Therefore

$$\sum_{\ell=1}^L \mathbb{E}|X_{n,\ell}|^p \leq \tilde{C}_p \sum_{\ell=1}^L \frac{n^p}{N_\ell^{p/2} m^{p\ell/2}} \leq \frac{\tilde{C}_p}{\left( \sum_{\ell=1}^L a_\ell \right)^{p/2}} \sum_{\ell=1}^L a_\ell^{p/2} \xrightarrow{n \rightarrow \infty} 0.$$

This completes the proof.  $\square$

### 3.3 Complexity analysis

As in the Monte Carlo case we can interpret Theorem 5 as follows. For a total error of order  $1/n^\alpha$  the computational effort necessary to run the multilevel Monte Carlo Euler method is given by the sequence of sample sizes specified by relation (20). The associated time complexity is given by:

$$\begin{aligned} C_{MMC} &= C \times \left( n^{2\alpha} + \sum_{\ell=1}^L N_\ell (m^\ell + m^{\ell-1}) \right) \quad \text{with } C > 0 \\ &= C \times \left( n^{2\alpha} + n^{2\alpha} \frac{(m^2 - 1)T}{m} \sum_{\ell=1}^L \frac{1}{a_\ell} \sum_{\ell=1}^L a_\ell \right). \end{aligned}$$

The minimum of this complexity is reached for the choice of weights  $a_\ell^* = 1$ ,  $\ell \in \{1, \dots, L\}$ , since the Cauchy-Schwartz inequality ensures that  $L^2 \leq \sum_{\ell=1}^L \frac{1}{a_\ell} \sum_{\ell=1}^L a_\ell$ , and the optimal complexity for the multilevel Monte Carlo Euler method is given by

$$C_{MMC} = C \times \left( n^{2\alpha} + n^{2\alpha} (\log n)^2 \frac{m^2 - 1}{m(\log m)^2} \right) = O(n^{2\alpha} (\log n)^2).$$

Note that this optimal choice  $a_\ell^* = 1$  corresponds to the sample sizes given by

$$N_\ell = \frac{(m - 1)T n^{2\alpha} \log n}{m^\ell \log m}, \quad \ell \in \{1, \dots, L\}.$$

Hence, our optimal choice is consistent with that proposed by Giles [8]. Nevertheless, unlike the parameters obtained by Giles [8] for the same setting (see relation (1)), our optimal choice of the sample sizes  $N_\ell$ ,  $\ell \in \{1, \dots, L\}$  does not depend on any given constant, since our approach is based on proving a central limit theorem and not on getting upper bounds for the variance. Note also that the optimal choice of the parameter  $m$  is obtained for  $m^* = 7$ . Otherwise, for the same error of order  $1/n^\alpha$  we have shown that the optimal complexity of a Monte Carlo method was given by

$$C_{MC} = C \times n^{2\alpha+1}$$

which is clearly larger than  $C_{MMC}$ . So we deduce that the multilevel method is more efficient.

## 4 Multilevel Monte Carlo and Asian options

The payoff of an Asian option is related to the integral of the asset price process. Computing the price of an Asian option requires the discretization of the integral. The purpose of this section is to apply Multilevel Monte Carlo to the approximation of the integral and to carry on a complexity analysis in this context. This will lead us to prove a stable law convergence theorem for the discretization error, which can be viewed as the analogue of our Theorem 4 above. Let  $S$  be the process on the stochastic basis  $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  satisfying

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t, \quad \text{with } t \in [0, T], \quad T > 0,$$

where  $\sigma$  and  $r$  are real constants, with  $\sigma > 0$  and  $(W_t)_{t \in [0, T]}$  is a standard Brownian motion on  $\mathcal{B}$ . The solution of the last equation is given by

$$S_t = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma W_t\right), \quad \text{where } S_0 > 0.$$

We set

$$I_T = \frac{1}{T} \int_0^T S_u du.$$

Let  $f$  be a given real valued function. Our aim will be to evaluate

$$e^{-rT} \mathbb{E} f(I_T).$$

In a financial setting, if  $f(x) = (x - K)_+$ , this last quantity is the price of an Asian call option with fixed strike  $K$ . In this case there is no explicit formula that gives the real price. So, the computation of this price, by a probabilistic method, requires a discretization of the integral  $I_T$ . There are several approximation schemes used in practice and one can consider either Riemann scheme or the trapezoidal scheme. We have

$$I_T^n = \frac{1}{n} \sum_{k=0}^{n-1} S_{k\delta}, \quad \text{and} \quad J_T^n = \frac{1}{n} \sum_{k=0}^{n-1} \frac{S_{k\delta} + S_{(k+1)\delta}}{2}, \quad \text{where } \delta = \frac{T}{n}.$$

We call the first approximation Riemann scheme because it is closely related to the Riemann approximation of the integral and the second the trapezoidal scheme because it is closely related to the trapezoidal approximation of the integral. We recall some results proved by Lapeyre and Temam [21] on the expansions for the strong and weak errors associated to both Riemann and trapezoidal schemes. Indeed, concerning the strong errors they prove that for  $p \geq 1$ , there exist  $K_p(T) > 0$  and  $\tilde{K}_p(T) > 0$  such that

$$\begin{aligned} \mathcal{PR1)} \quad & \left( \mathbb{E} \left( \sup_{t \in [0, T]} |I_t^n - I_t|^{2p} \right) \right)^{1/2p} \leq \frac{K_p(T)}{n}, \\ \mathcal{PT1)} \quad & \left( \mathbb{E} \left( \sup_{t \in [0, T]} |J_t^n - I_t|^{2p} \right) \right)^{1/2p} \leq \frac{\tilde{K}_p(T)}{n}. \end{aligned}$$

Hence, it is obvious that both schemes are of second order. Concerning the weak errors they prove that for any  $\mathbb{R}$ -valued function  $f$  satisfying condition  $(\mathcal{H}_f)$ , if  $\mathbb{P}(I_T \notin \mathcal{D}_f) = 0$ , where  $\mathcal{D}_f := \{x \in \mathbb{R}^d; f \text{ is differentiable at } x\}$ , then there exist real constants  $C_f^I$  and  $C_f^J$  such that

$$\begin{aligned} \mathcal{PR2)} \quad & \lim_{n \rightarrow \infty} n \left( \mathbb{E} f(I_T^n) - \mathbb{E} f(I_T) \right) = C_f^I, \\ \mathcal{PT2)} \quad & \lim_{n \rightarrow \infty} n \left( \mathbb{E} f(J_T^n) - \mathbb{E} f(I_T) \right) = C_f^J. \end{aligned}$$

In order to obtain central limit theorems for the multilevel Monte Carlo method associated with both Riemann and Trapezoidal schemes, we study the asymptotic behavior of the distribution errors as in the previous Euler scheme case. We establish two stable convergence theorems for each scheme.

## 4.1 Stable convergence of the Riemann scheme error

The Riemann approximation of the process is given by

$$I_t^n = \frac{1}{T} \int_0^t S_{\eta_n(u)} du = \frac{1}{n} \sum_{k=0}^{[t/\delta]-1} S_{k\delta} + \frac{t - \eta_n(t)}{T} S_{\eta_n(t)}$$

with  $\eta_n(t) = [t/\delta]\delta$ . One have to study the distribution of the error process  $I^{mn} - I^n$ .

**Theorem 6** *We have the following result*

$$\frac{mn}{\sqrt{m^2 - 1}}(I^{mn} - I^n) \Rightarrow^{stably} \xi$$

where  $\xi$  is the process defined by

$$\xi_t := \sqrt{\frac{m-1}{m+1}} \frac{S_t - S_0}{2} + \frac{1}{2\sqrt{3}} \int_0^t \sigma S_u dB_u,$$

where  $B$  is a standard Brownian motion on an extension  $\hat{\mathcal{B}}$  of  $\mathcal{B}$ , which is independent of  $W$ .

**Proof** The error,  $\mathcal{E}_t^{mn,n}$ , is given by

$$\mathcal{E}_t^{mn,n} := I_t^{mn} - I_t^n = \frac{1}{T} \int_0^t (S_{\eta_{mn}(s)} - S_{\eta_n(s)}) ds.$$

Noting that the integrand vanishes on the interval  $[\eta_n(s), \eta_n(s) + \frac{1}{mn}]$ , this error can be written as follows

$$\begin{aligned} \mathcal{E}_t^{mn,n} &= \frac{1}{T} \int_0^t (S_{\eta_{mn}(s)} - S_{\eta_n(s)}) \mathbf{1}_{\{\eta_n(s) + \frac{1}{mn} \leq s < \eta_n(s) + \frac{1}{n}\}} ds + R_t^{mn,n} \\ &= \frac{1}{mn} \sum_{k=0}^{[t/\delta]-1} \sum_{\ell=1}^{m-1} (S_{(mk+\ell)\delta/m} - S_{k\delta}) + R_t^{mn,n}, \end{aligned}$$

where  $R_t^{mn,n} = \frac{1}{T} \int_{\eta_n(t)}^t (S_{\eta_{mn}(s)} - S_{\eta_n(s)}) ds$ . Now, using the dynamic of  $S_t$  we get

$$\begin{aligned} \mathcal{E}_t^{mn,n} &= \frac{1}{mn} \sum_{k=0}^{[t/\delta]-1} \sum_{\ell=1}^{m-1} \int_{k\delta}^{(mk+\ell)\delta/m} r S_u du + \frac{1}{mn} \sum_{k=0}^{[t/\delta]-1} \sum_{\ell=1}^{m-1} \int_{k\delta}^{(mk+\ell)\delta/m} \sigma S_u dW_u + R_t^{mn,n} \\ &= \frac{1}{mn} \sum_{k=0}^{[t/\delta]-1} \sum_{\ell=1}^{m-1} \int_{k\delta}^{(k+1)\delta} S_u \mathbf{1}_{\{k\delta \leq u < (mk+\ell)\delta/m\}} (r du + \sigma dW_u) + R_t^{mn,n} \\ &= \frac{1}{mn} \int_0^{\eta_n(t)} r S_u \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) du + \frac{1}{mn} \int_0^{\eta_n(t)} \sigma S_u \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) dW_u + R_t^{mn,n} \end{aligned}$$

with the digital function defined, for  $\ell \in \{1, \dots, m-1\}$ , by

$$d_\ell^{mn,n}(u) := \mathbf{1}_{\{\eta_n(u) \leq u < \eta_n(u) + \ell\delta/m\}}.$$

Hence we get

$$mn\mathcal{E}_t^{mn,n} = \int_0^t rS_u dD_u^{mn,n} + \int_0^t \sigma S_u dM_u^{mn,n} + mnR_t^{mn,n}$$

with the martingale integrand

$$M_t^{mn,n} := \int_0^{\eta_n(t)} \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) dW_u,$$

and a drift term with bounded variation

$$D_t^{mn,n} := \int_0^{\eta_n(t)} \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) du.$$

To study the convergence of the martingale, we compute its quadratic variation

$$\langle M^{mn,n} \rangle_t = \int_0^{\eta_n(t)} \sum_{\ell=1}^{m-1} (d_\ell^{mn,n}(u))^2 du + 2 \int_0^{\eta_n(t)} \sum_{1 \leq \ell < \ell' \leq m-1} d_\ell^{mn,n}(u) d_{\ell'}^{mn,n}(u) du \quad (23)$$

For the first term, we note

$$\sum_{\ell=1}^{m-1} (d_\ell^{mn,n}(u))^2 = \sum_{\ell=1}^{m-1} \mathbf{1}_{\{\eta_n(u) \leq u < \eta_n(u) + \ell\delta/m\}}.$$

Concerning the second integral, since for  $1 \leq \ell < \ell' \leq m-1$  we have

$$\eta_n(u) \leq \eta_n(u) + \ell\delta/m \leq \eta_n(u) + \ell'\delta/m \leq \eta_n(u) + \delta,$$

the expansion of  $d_\ell^{mn,n}(u) d_{\ell'}^{mn,n}(u) = d_\ell^{mn,n}(u)$ . Therefore, coming back to the bracket (23), we get after computation

$$\langle M^{mn,n} \rangle_t = \frac{[t/\delta]\delta}{m} \sum_{\ell=1}^{m-1} \ell + \frac{2[t/\delta]\delta}{m} \sum_{1 \leq \ell < \ell' \leq m-1} \ell \xrightarrow{n \rightarrow \infty} \frac{(m-1)(2m-1)}{6} t.$$

Furthermore, by simple computation we get

$$\langle M^{mn,n}, W \rangle_t = \int_0^{\eta_n(t)} \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) du = \frac{[t/\delta]\delta}{m} \sum_{\ell=1}^{m-1} \ell = \frac{m-1}{2} [t/\delta]\delta \xrightarrow{n \rightarrow \infty} \frac{m-1}{2} t.$$

Besides, using this last computation it is easy to check that

$$\sup_{0 \leq t \leq T} \left| D_t^{mn,n} - \frac{m-1}{2} t \right| = \frac{m-1}{2} \sup_{0 \leq t \leq T} \left| [t/\delta]\delta - t \right| \xrightarrow{n \rightarrow \infty} 0.$$



By virtue of Theorem 2-1 of Jacod [14] we obtain the stable convergence of

$$rD_t^{mn,n} + \sigma M_t^{mn,n} \Rightarrow^{stably} \frac{m-1}{2}rt + \frac{m-1}{2}\sigma W_t + \sqrt{\frac{m^2-1}{12}}\sigma B_t$$

where  $(B_t)_{t \geq 0}$  is a Brownian motion independent of  $(W_t)_{t \geq 0}$ . Moreover, according to the above computations, it is easy to check the tightness of  $\langle M^{mn,n} \rangle_T + \int_0^T d|D_t^{mn,n}|$  and thanks to Lemma 1 and Theorem 1 we get

$$\int_0^t rS_u dD_u^{mn,n} + \int_0^t \sigma S_u dM_u^{mn,n} \Rightarrow^{stably} (m-1)\frac{S_t - S_0}{2} + \sqrt{\frac{m^2-1}{12}} \int_0^t \sigma S_u dB_u.$$

Now, it remains to prove the convergence of  $\sup_{0 \leq t \leq T} |mnR_t^{mn,n}|$  in probability to zero. This rest term is bounded up to a constant factor by  $\sup_{0 \leq t \leq T} |S_{\eta_{mn}(t)} - S_{\eta_n(t)}|$ . Finally, the proof is completed using the hölder regularity of the process  $S$ . □

The above subsection is devoted to the study of the trapezoidal scheme error.

## 4.2 Stable convergence of the trapezoidal scheme error

The trapezoidal approximation of the process is given by

$$J_t^n = \frac{1}{T} \int_0^t \frac{S_{\eta_n(u)} + S_{(\eta_n(u)+\delta) \wedge t}}{2} du = \frac{1}{n} \sum_{k=0}^{[t/\delta]-1} \frac{S_{k\delta} + S_{(k+1)\delta}}{2} + (t - \eta_n(t)) \frac{S_{\eta_n(t)} + S_t}{2T}$$

with  $\eta_n(t) = [t/\delta]\delta$ . One have to study the error process given by  $J^{mn} - J^n$ .

**Theorem 7** *We have the following result*

$$\frac{mn}{\sqrt{m^2-1}}(J^{mn} - J^n) \Rightarrow^{stably} \chi$$

where  $\chi$  is the process defined by

$$\chi_t := \frac{1}{2\sqrt{3}} \int_0^t \sigma S_u dB_u,$$

where  $B$  is a standard Brownian motion on an extension  $\hat{\mathcal{B}}$  of  $\mathcal{B}$ , which is independent of  $W$ .

**Remark** The process  $\chi$  above is the same limit process given in Theorem 4.1 of Kebaier [16]. In fact, he proves that

$$m^\ell(J^{m^\ell} - J) \Rightarrow^{stably} \chi, \quad \text{as } \ell \rightarrow \infty,$$

which is not sufficient to prove our Theorem 9 below, since the multilevel Monte Carlo method involves the error process  $J^{m^\ell} - J^{m^{\ell-1}}$  rather than  $J^{m^\ell} - J$ .

**Proof** Considering the trapezoidal scheme, for the fine time discretization step  $\delta/m$ , we can write it as follows

$$\begin{aligned} J_t^{mn} &= \frac{1}{2T} \int_0^{\eta_n(t)} (S_{\eta_{mn}(u)} + S_{(\eta_{mn}(u)+\delta/m)}) du + \frac{1}{2T} \int_{\eta_n(t)}^t (S_{\eta_{mn}(u)} + S_{(\eta_{mn}(u)+\delta/m) \wedge t}) du \\ &= \frac{1}{2mn} \sum_{\ell=0}^{m-1} \sum_{k=0}^{[t/\delta]-1} (S_{(mk+\ell)\delta/m} + S_{(mk+\ell+1)\delta/m}) + \frac{1}{2T} \int_{\eta_n(t)}^t (S_{\eta_{mn}(u)} + S_{(\eta_{mn}(u)+\delta/m) \wedge t}) du. \end{aligned}$$

The first term in the right hand side, can be arranged as follows

$$\frac{1}{2mn} \sum_{k=0}^{[t/\delta]-1} (S_{k\delta} + S_{(k+1)\delta}) + \frac{1}{mn} \sum_{\ell=1}^{m-1} \sum_{k=0}^{[t/\delta]-1} S_{(mk+\ell)\delta/m}.$$

So that, the error,  $\mathcal{E}_t^{mn,n}$ , can be arranged as follows

$$\mathcal{E}_t^{mn,n} := J_t^{mn} - J_t^n = \frac{1-m}{2mn} \sum_{k=0}^{[t/\delta]-1} (S_{k\delta} + S_{(k+1)\delta}) + \frac{1}{mn} \sum_{\ell=1}^{m-1} \sum_{k=0}^{[t/\delta]-1} S_{(mk+\ell)\delta/m} + R_t^{mn,n},$$

with

$$R_t^{mn,n} = \frac{1}{2T} \int_{\eta_n(t)}^t (S_{\eta_{mn}(u)} + S_{(\eta_{mn}(u)+\delta/m) \wedge t} - S_{\eta_n(u)} - S_{(\eta_n(u)+\delta) \wedge t}) du.$$

Furthermore, we rewrite the error as

$$\mathcal{E}_t^{mn,n} = -\frac{1}{2mn} \sum_{k=0}^{[t/\delta]-1} \sum_{\ell=1}^{m-1} ((S_{(k+1)\delta} - S_{(mk+\ell)\delta/m}) - (S_{(mk+\ell)\delta/m} - S_{k\delta})) + R_t^{mn,n}.$$

Now, using the dynamic of  $S_t$  we get

$$\begin{aligned} \mathcal{E}_t^{mn,n} &= -\frac{1}{2mn} \sum_{k=0}^{[t/\delta]-1} \sum_{\ell=1}^{m-1} \int_{k\delta}^{(k+1)\delta} r S_u (\mathbf{1}_{\{(mk+\ell)\delta/m \leq u < (k+1)\delta\}} - \mathbf{1}_{\{k\delta \leq u < (mk+\ell)\delta/m\}}) du \\ &\quad - \frac{1}{2mn} \sum_{k=0}^{[t/\delta]-1} \sum_{\ell=1}^{m-1} \int_{k\delta}^{(k+1)\delta} \sigma S_u (\mathbf{1}_{\{(mk+\ell)\delta/m \leq u < (k+1)\delta\}} - \mathbf{1}_{\{k\delta \leq u < (mk+\ell)\delta/m\}}) dW_u + R_t^{mn,n} \\ &= -\frac{1}{2mn} \int_0^{\eta_n(t)} r S_u \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) du - \frac{1}{2mn} \int_0^{\eta_n(t)} \sigma S_u \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) dW_u + R_t^{mn,n} \end{aligned}$$

where the digital function defined, for  $\ell \in \{1, \dots, m-1\}$ , by

$$d_\ell^{mn,n}(u) := \mathbf{1}_{\{\eta_n(u)+\ell\delta/m \leq u < \eta_n(u)+\delta\}} - \mathbf{1}_{\{\eta_n(u) \leq u < \eta_n(u)+\ell\delta/m\}}.$$

Hence, we get

$$mn\mathcal{E}_t^{mn,n} = \int_0^t r S_u dD_u^{mn,n} + \int_0^t \sigma S_u dM_u^{mn,n} + mnR_t^{mn,n}$$

with the martingale integrand

$$M_t^{mn,n} := -\frac{1}{2} \int_0^{\eta_n(t)} \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) dW_u,$$

and a drift term

$$D_t^{mn,n} := -\frac{1}{2} \int_0^{\eta_n(t)} \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) du.$$

To study the convergence of the martingale, we compute its quadratic variation

$$4\langle M^{mn,n} \rangle_t = \int_0^{\eta_n(t)} \sum_{\ell=1}^{m-1} (d_\ell^{mn,n}(u))^2 du + 2 \int_0^{\eta_n(t)} \sum_{1 \leq \ell < \ell' \leq m-1} d_\ell^{mn,n}(u) d_{\ell'}^{mn,n}(u) du \quad (24)$$

For the first term, we note that

$$\begin{aligned} \sum_{\ell=1}^{m-1} (d_\ell^{mn,n}(u))^2 &= \sum_{\ell=1}^{m-1} \mathbf{1}_{\{\eta_n(u) + \ell\delta/m \leq u < \eta_n(u) + \delta\}} + \mathbf{1}_{\{\eta_n(u) \leq u < \eta_n(u) + \ell\delta/m\}} \\ &= \sum_{\ell=1}^{m-1} \mathbf{1}_{\{\eta_n(u) \leq u < \eta_n(u) + \delta\}} = (m-1). \end{aligned}$$

Concerning the second integral, since for  $1 \leq \ell < \ell' \leq m-1$  we have

$$\eta_n(u) \leq \eta_n(u) + \ell\delta/m \leq \eta_n(u) + \ell'\delta/m \leq \eta_n(u) + \delta,$$

the expansion of  $d_\ell^{mn,n}(u) d_{\ell'}^{mn,n}(u)$  is equal to

$$\mathbf{1}_{\{\eta_n(u) + \ell'\delta/m \leq u < \eta_n(u) + \delta\}} - \mathbf{1}_{\{\eta_n(u) + \ell\delta/m \leq u < \eta_n(u) + \ell'\delta/m\}} + \mathbf{1}_{\{\eta_n(u) \leq u < \eta_n(u) + \ell\delta/m\}}$$

that we rewrite as  $1 - 2 \times \mathbf{1}_{\{\eta_n(u) + \ell\delta/m \leq u < \eta_n(u) + \ell'\delta/m\}}$ . Coming back to the bracket (24), we get after computation

$$4\langle M^{mn,n} \rangle_t = (m-1)^2 t - \frac{4[t/\delta]\delta}{m} \sum_{1 \leq \ell < \ell' \leq m-1} (\ell' - \ell) \xrightarrow{n \rightarrow \infty} \frac{m^2 - 1}{3} t.$$

Furthermore, by simple computation we get

$$-2\langle M^{mn,n}, W \rangle_t = \int_0^{\eta_n(t)} \sum_{\ell=1}^{m-1} d_\ell^{mn,n}(u) du = \frac{[t/\delta]\delta}{m} \sum_{\ell=1}^{m-1} (m - \ell) - \frac{[t/\delta]\delta}{m} \sum_{\ell=1}^{m-1} \ell = 0.$$

Finally, We can proceed analogously to the Riemann case to achieve the proof.  $\square$

We can now formulate our main results for both Riemann and trapezoidal schemes.

### 4.3 Central Limit Theorems

In the same way as in Euler scheme frame, we consider a real sequence  $(a_\ell)_{\ell \geq 1}$  of positive terms satisfying

$$(\mathcal{W}) \quad \lim_{n \rightarrow \infty} \sum_{\ell=1}^L a_\ell = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{1}{\left(\sum_{\ell=1}^L a_\ell\right)^{p/2}} \sum_{\ell=1}^L a_\ell^{p/2} = 0, \quad \text{for } p > 2. \quad (25)$$

Let us assume that the sample sizes  $N_\ell$ , for  $\ell \in \{1, \dots, L\}$ , for the multilevel Monte Carlo method, have the following form

$$N_\ell = \frac{n^2(m^2 - 1)}{m^{2\ell} a_\ell} \sum_{\ell=1}^L a_\ell, \quad \ell \in \{1, \dots, L\} \quad \text{and} \quad L = \frac{\log n}{\log m}. \quad (26)$$

#### 4.3.1 Riemannian Scheme

Now, we consider the Riemann scheme

$$\mathbb{E}(f(I_T^n)) = \mathbb{E}(f(I_T^1)) + \sum_{\ell=1}^L \mathbb{E}\left(f(I_T^{m^\ell}) - f(I_T^{m^{\ell-1}})\right). \quad (27)$$

It is worth to note that  $f(I_T^1)$  is deterministic equal to  $f(s_0)$ . Hence, the multilevel method in this case can be written as

$$Q_n = f(s_0) + \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(f(I_{T,k}^{m^\ell}) - f(I_{T,k}^{m^{\ell-1}})\right). \quad (28)$$

We can now state the Central limit theorem in this setting.

**Theorem 8** *Let  $f$  be a  $\mathbb{R}$ -valued function satisfying condition  $(\mathcal{H}_f)$  and such that  $\mathbb{P}(I_T \notin \mathcal{D}_f) = 0$ , where  $\mathcal{D}_f := \{x \in \mathbb{R}^d; f \text{ is differentiable at } x\}$ . We have*

$$n(Q_n - \mathbb{E}(f(I_T))) \Rightarrow \mathcal{N}(C_f^I, \sigma^2)$$

where  $\sigma^2 = \tilde{Var}(f'(I_T)\xi_T)$  and  $C_f^I$  is given by property  $\mathcal{PR2}$ ). Here  $\xi$  is the limit process in Theorem 6.

**Proof :** Combining relations (27) and (28) we obtain

$$Q_n - \mathbb{E}(f(I_T)) = \hat{Q}_n + \mathbb{E}(f(I_T^n)) - \mathbb{E}(f(I_T)),$$

where

$$\hat{Q}_n = \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(f(I_{T,k}^{m^\ell}) - f(I_{T,k}^{m^{\ell-1}}) - \mathbb{E}\left(f(I_T^{m^\ell}) - f(I_T^{m^{\ell-1}})\right)\right)$$

Using assumption  $\mathcal{PR2}$ ), we obviously obtain the term  $C_f^I$  in the limit. So, we have only to establish

$$n\hat{Q}_n \Rightarrow \mathcal{N}\left(0, \tilde{Var}(f'(I_T)\xi_T)\right).$$

To do so, we plan to use the Lindeberg Feller theorem [10] with Lyapunov condition. More precisely, We set

$$X_{n,\ell} := \frac{n}{N_\ell} \sum_{k=1}^{N_\ell} Z_{T,k}^{m^\ell, m^{\ell-1}} \text{ and } Z_{T,k}^{m^\ell, m^{\ell-1}} := f(I_{T,k}^{m^\ell}) - f(I_{T,k}^{m^{\ell-1}}) - \mathbb{E} \left( f(I_{T,K}^{m^\ell}) - f(I_{T,K}^{m^{\ell-1}}) \right),$$

and we have only to check the following conditions :

- $\lim_{n \rightarrow \infty} \sum_{\ell=1}^L \mathbb{E}(X_{n,\ell})^2 = \tilde{Var}(f'(I_T)\xi_T)$
- (Lyapunov condition) there exists  $p > 2$  such that  $\lim_{n \rightarrow \infty} \sum_{\ell=1}^L \mathbb{E}|X_{n,\ell}|^p = 0$ .

For the first one, we have

$$\begin{aligned} \sum_{\ell=1}^L \mathbb{E}(X_{n,\ell})^2 &= \sum_{\ell=1}^L Var(X_{n,\ell}) = \sum_{\ell=1}^L \frac{n^2}{N_\ell} Var \left( Z_{T,1}^{m^\ell, m^{\ell-1}} \right) \\ &= \frac{1}{\sum_{\ell=1}^L a_\ell} \sum_{\ell=1}^L a_\ell \frac{m^{2\ell}}{m^2 - 1} Var \left( Z_{T,1}^{m^\ell, m^{\ell-1}} \right). \end{aligned} \quad (29)$$

Otherwise, since  $\mathbb{P}(I_T \notin \mathcal{D}_f) = 0$ , applying the Taylor expansion theorem twice we get

$$\begin{aligned} f(I_T^{m^\ell}) - f(I_T^{m^{\ell-1}}) &= f'(I_T)(I_T^{m^\ell} - I_T^{m^{\ell-1}}) + \\ &\quad (I_T^{m^\ell} - I_T)\varepsilon(I_T, I_T^{m^\ell} - I_T) - (I_T^{m^{\ell-1}} - I_T)\varepsilon(I_T, I_T^{m^{\ell-1}} - I_T). \end{aligned}$$

with  $\varepsilon(I_T, I_T^{m^\ell} - I_T) \xrightarrow[\ell \rightarrow \infty]{\mathbb{P}} 0$  and  $\varepsilon(I_T, I_T^{m^{\ell-1}} - I_T) \xrightarrow[\ell \rightarrow \infty]{\mathbb{P}} 0$ . By property  $\mathcal{PR1}$ , we get the tightness of

$\frac{m^\ell}{\sqrt{m^2-1}}(I_T^{m^\ell} - I_T)$  and  $\frac{m^\ell}{\sqrt{m^2-1}}(I_T^{m^{\ell-1}} - I_T)$  and we deduce

$$\frac{m^\ell}{\sqrt{m^2-1}} \left( (I_T^{m^\ell} - I_T)\varepsilon(I_T, I_T^{m^\ell} - I_T) - (I_T^{m^{\ell-1}} - I_T)\varepsilon(I_T, I_T^{m^{\ell-1}} - I_T) \right) \xrightarrow[\ell \rightarrow \infty]{\mathbb{P}} 0.$$

So, according to lemma 1 and Theorem 6 and since  $f'(I_T^{m^{\ell-1}}) \xrightarrow[\ell \rightarrow \infty]{\mathbb{P}} f'(I_T)$  we conclude that

$$\frac{m^\ell}{\sqrt{m^2-1}} \left( f(I_T^{m^\ell}) - f(I_T^{m^{\ell-1}}) \right) \Rightarrow^{stably} f'(I_T)\xi_T, \text{ as } \ell \rightarrow \infty. \quad (30)$$

Now, using  $(\mathcal{H}_f)$  it follows from property  $\mathcal{PR1}$  that

$$\forall \varepsilon > 0, \quad \sup_{\ell} \mathbb{E} \left| \frac{m^\ell}{\sqrt{m^2-1}} \left( f(I_T^{m^\ell}) - f(I_T^{m^{\ell-1}}) \right) \right|^{2+\varepsilon} < \infty.$$

We deduce using (30) that

$$\mathbb{E} \left( \frac{m^\ell}{\sqrt{m^2-1}} \left( f(I_T^{m^\ell}) - f(I_T^{m^{\ell-1}}) \right) \right)^k \rightarrow \mathbb{E} \left( f'(I_T)\xi_T \right)^k < \infty \text{ with } k \in \{1, 2\}.$$

Consequently,

$$\frac{m^{2\ell}}{m^2 - 1} \text{Var}(Z_{T,1}^{m^\ell, m^{\ell-1}}) \longrightarrow \tilde{\text{Var}}(f'(I_T)\xi_T) < \infty.$$

Combining this last convergence with relation (29), we obtain the first condition a using Toeplitz lemma. Concerning the second one, by Burkholder's inequality and elementary computations, we get for  $p > 2$

$$\mathbb{E}|X_{n,\ell}|^p = \frac{n^p}{N_\ell^p} \mathbb{E} \left| \sum_{\ell=1}^{N_\ell} Z_{T,1}^{m^\ell, m^{\ell-1}} \right|^p \leq C_p \frac{n^p}{N_\ell^{p/2}} \mathbb{E} \left| Z_{T,1}^{m^\ell, m^{\ell-1}} \right|^p,$$

where  $C_p$  is a numerical constant that depends on  $p$  only. Otherwise, property  $\mathcal{PR1}$ ) ensures the existence of a constant  $K_p > 0$  such that

$$\mathbb{E} \left| Z_{T,1}^{m^\ell, m^{\ell-1}} \right|^p \leq \frac{K_p}{m^{p\ell}}.$$

Therefore,

$$\sum_{\ell=1}^L \mathbb{E} |X_{n,\ell}|^p \leq \tilde{C}_p \sum_{\ell=1}^L \frac{n^p}{N_\ell^{p/2} m^{p\ell}} \leq \frac{\tilde{C}_p}{\left(\sum_{\ell=1}^L a_\ell\right)^{p/2}} \sum_{\ell=1}^L a_\ell^{p/2} \xrightarrow{n \rightarrow \infty} 0.$$

This completes the proof.  $\square$

#### 4.3.2 Trapezoidal Scheme

Now, we consider the trapezoidal scheme.

$$\mathbb{E}(f(J_T^n)) = \mathbb{E}(f(J_T^1)) + \sum_{\ell=1}^L \mathbb{E} \left( f(J_T^{m^\ell}) - f(J_T^{m^{\ell-1}}) \right).$$

Hence, the multilevel method in this case can be written as

$$Q_n = \frac{1}{N_0} \sum_{k=1}^{N_0} f \left( \frac{S_0 + S_{T,k}}{2} \right) + \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left( f(I_{T,k}^{m^\ell}) - f(I_{T,k}^{m^{\ell-1}}) \right).$$

In the following we consider the same real sequence  $(a_\ell)_{\ell \geq 1}$  of positive terms given by relation (25) and the sequence of sample sizes  $(N_\ell)_{0 \leq \ell \leq L}$  given by relation (26). We can now state the Central limit theorem for the trapezoidal scheme.

**Theorem 9** *Let  $f$  be a  $\mathbb{R}$ -valued function satisfying condition  $(\mathcal{H}_f)$  and such that  $\mathbb{P}(I_T \notin \mathcal{D}_f) = 0$ , where  $\mathcal{D}_f := \{x \in \mathbb{R}^d; f \text{ is differentiable at } x\}$ . We have*

$$n(Q_n - \mathbb{E}(f(I_T))) \Rightarrow \mathcal{N}(C_f^J, \sigma^2)$$

where

$$\sigma^2 = \text{Var} \left( f \left( \frac{S_0 + S_T}{2} \right) \right) + \tilde{\text{Var}}(f'(I_T)\chi_T)$$

and  $C_f^J$  is given by property  $\mathcal{PT2}$ ). Here,  $\chi$  is the limit process in Theorem 7.

**Proof:** We can write

$$Q_n - \mathbb{E}(f(I_T)) = \hat{Q}_n^1 + \hat{Q}_n^2 + \mathbb{E}(f(J_T^n)) - \mathbb{E}(f(I_T)),$$

where

$$\begin{aligned}\hat{Q}_n^1 &= \frac{1}{N_0} \sum_{k=1}^{N_0} \left( f(J_{T,k}^{m^0}) - \mathbb{E}\left(f(J_T^{m^0})\right) \right) \\ \hat{Q}_n^2 &= \sum_{\ell=1}^L \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left( f(J_{T,k}^{m^\ell}) - f(J_{T,k}^{m^{\ell-1}}) - \mathbb{E}\left(f(J_T^{m^\ell}) - f(J_T^{m^{\ell-1}})\right) \right).\end{aligned}$$

Using assumption  $\mathcal{PT2}$ ) we obviously obtain the term  $C_f^J$  in the limit. Afterward, one can apply the classical central limit theorem for the quantity  $\hat{Q}_n^1$  with  $N_0 = n^2$  to get

$$n\hat{Q}_n^1 \Rightarrow \mathcal{N}\left(0, \text{Var}(f(J_T^{m^0}))\right).$$

On the other hand, the convergence of  $n\hat{Q}_n^2$  is obtained by following the proof steps of the Central Limit Theorem for the Riemann scheme, Theorem 8. Using this approach, we have only to use respectively property  $\mathcal{PT1}$ ) and Theorem 7 instead of  $\mathcal{PR1}$ ) and Theorem 6. Hence, we obtain the following convergence

$$n\hat{Q}_n^2 \Rightarrow \mathcal{N}\left(0, \tilde{\text{Var}}(f'(I_T)\chi_T)\right).$$

This completes the proof. □

## 4.4 The complexity

The following complexity analysis stands valid for any second order discretization scheme. In particular, it remains valid for both Riemann and trapezoidal schemes. As in the Monte Carlo case we can interpret Theorems 8 and 9 as follows. For a total error of order  $1/n$  the computational effort necessary to run the multilevel algorithm applied to the Riemann or trapezoidal scheme, with step numbers  $m^\ell$ ,  $(m, \ell) \in \mathbb{N} \setminus \{0, 1\} \times \{1, \dots, L\}$ , corresponds to the sequence of sample sizes  $(N_\ell)_{0 \leq \ell \leq L}$  given by relation (26). Consequently, the time complexity in the multilevel Monte Carlo method for these second order schemes is given by

$$\begin{aligned}C_{MMC} &= C \times \sum_{\ell=1}^L N_\ell (m^\ell + m^{\ell-1}) \quad \text{with } C > 0 \\ &= C \times \frac{(m+1)^2(m-1)}{m} n^2 \sum_{\ell=1}^L \frac{1}{m^\ell a_\ell} \sum_{\ell=1}^L a_\ell.\end{aligned}$$

The minimum of this complexity is reached for the choice of weights  $a_\ell^* = m^{-\ell/2}$ ,  $\ell \in \{1, \dots, L\}$ , since the Cauchy-Schwartz inequality ensures that  $\left(\sum_{\ell=1}^L m^{-\ell/2}\right)^2 \leq \sum_{\ell=1}^L \frac{1}{m^\ell a_\ell} \sum_{\ell=1}^L a_\ell$ , and

the optimal complexity for the multilevel Monte Carlo method for this choice is given by

$$C_{MMC}^{a_\ell^*} = C \times \frac{(m+1)^2(m-1)}{m} n^2 \left( \sum_{\ell=1}^L m^{-\ell/2} \right)^2 = O(n^2).$$

Note that this optimal choice  $a_\ell^* = m^{-\ell/2}$  corresponds to the sample size

$$N_\ell = \frac{m^2 - 1}{m^{3\ell/2}(1 - \sqrt{m})} n^2 \left( 1 - \frac{1}{\sqrt{n}} \right) \quad (31)$$

of the  $\ell^{th}$  level in the multilevel algorithm, which is consistent with the complexity analysis given in Giles [8]. More precisely, by taking  $\beta = 2$  in Theorem 3.1 of [8] we recover the same complexity as well as the same order of sample sizes  $(N_\ell)_{0 \leq \ell \leq L}$  (see also relation (2)). However, this optimal choice  $a_\ell^*$ , leading to the complexity  $n^2$ , does not satisfy condition  $(\mathcal{W})$  and even the Lyapunov condition. Hence, there is no reason that the Central limit theorem holds. Actually, Giles's analysis is based on the control of the variance, whereas, for the same framework, to obtain the central limit theorem, we need in addition a Lyapunov type condition that controls a moment of order greater than 2. As a consequence for applications, we recommend not to use the multilevel Monte Carlo method associated to second order schemes with the sample size  $N_\ell$  given by relation (31), since the central limit theorem, essential when we use Monte Carlo methods, does not hold.

So, how to choose this sequence  $(N_\ell)_{0 \leq \ell \leq L}$  in an optimal way and such that the central limit theorem still holds? We shall exhibit three sequences  $(a_\ell)_{1 \leq \ell \leq L}$  satisfying our condition  $(\mathcal{W})$  and reducing significantly the complexity and for which the complexity is explicit. In the following, we fix  $N_0 = n^2$  and for all  $\ell \in \{1, \dots, L\}$  we have

a) the choice  $a_{\ell,1} = 1$ , corresponds to the sample size  $N_{\ell,1} = \frac{m^2 - 1}{m^{2\ell}} n^2 L$ . This leads to a complexity  $C_{MMC}^{a_{\ell,1}} = C \times \left( n^2 + \frac{(m+1)^2}{m \log m} n^2 \log n \right) = O(n^2 \log n)$ . In this case, the optimal choice of the parameter  $m$  is equal to 4.

b) For  $a_{\ell,2} = 1/\ell$ , we get  $N_{\ell,2} = \frac{(m^2 - 1)\ell}{m^{2\ell}} n^2 \sum_{\ell=1}^L \frac{1}{\ell}$ . This leads to a complexity

$$\begin{aligned} C_{MMC}^{a_{\ell,2}} &= C \times \left( n^2 + \frac{(m+1)^2(m-1)}{m} n^2 \sum_{\ell=1}^L \frac{\ell}{m^\ell} \sum_{\ell=1}^L \frac{1}{\ell} \right) \\ &\sim C \times \frac{(m+1)^2}{m-1} n^2 \log \log n = O(n^2 \log \log n) \end{aligned}$$

and the optimal choice of the parameter  $m$  is equal to 3.



c) For  $a_{\ell,3} = 1/(\ell \log \ell)$ , we get  $N_{\ell,3} = \frac{(m^2 - 1)\ell \log \ell}{m^{2\ell}} n^2 \sum_{\ell=1}^L \frac{1}{\ell \log \ell}$  and a complexity

$$\begin{aligned} C_{MMC}^{a_{\ell,3}} &= C \times \left( n^2 + \frac{(m+1)^2(m-1)}{m} n^2 \sum_{\ell=1}^L \frac{\ell \log \ell}{m^\ell} \sum_{\ell=1}^L \frac{1}{\ell \log \ell} \right) \\ &\sim C \times \frac{(m+1)^2(m-1)}{m} n^2 \sum_{\ell=1}^{\infty} \frac{\ell \log \ell}{m^\ell} n^2 \log \log \log n = O(n^2 \log \log \log n). \end{aligned}$$

In this last case, the factor depending on  $m$ , in the above complexity, can be interpreted as  $\frac{(m+1)^2}{m} \mathbb{E}(G_m \log(G_m))$ , where  $G_m \stackrel{\text{law}}{=} \text{Geometric}(1 - 1/m)$ . So, a simple Monte Carlo approximation yields the optimal choice of the parameter  $m$  which is equal to 5.

Through these examples, we note that the central limit theorem is conserved and the complexity can be very close to the order  $n^2$  without reaching it. In the other hand, for the same error of order  $1/n$  we have shown that the optimal complexity of a Monte Carlo method was given by:

$$C_{MC} = C \times n^3$$

which is clearly larger than  $C_{MMC}$ . So we deduce that the multilevel method is more efficient.

## 5 Conclusion

The multilevel Monte Carlo algorithm is a method that can be used in a general framework: as soon as we use a discretization scheme in order to compute quantities such as  $\mathbb{E}f(X_t, 0 \leq t \leq T)$ , we can implement the statistical multilevel algorithm. And this is worth because it is more efficient than a classic Monte Carlo method.

## References

- [1] D. J. Aldous and G. K. Eagleson. On mixing and stability of limit theorems. *Ann. Probability*, 6(2):325–331, 1978.
- [2] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Related Fields*, 104(1):43–60, 1996.
- [3] J. Creutzig, S. Dereich, T. Müller-Gronbach, and K. Ritter. Infinite-dimensional quadrature and approximation of distributions. *Found. Comput. Math.*, 9(4):391–429, 2009.
- [4] S. Dereich. Multilevel Monte Carlo algorithms for Lévy-driven SDEs with Gaussian correction. *Ann. Appl. Probab.*, 21(1):283–311, 2011.
- [5] D. Duffie and P. Glynn. Efficient Monte Carlo simulation of security prices. *Ann. Appl. Probab.*, 5(4):897–905, 1995.

- [6] O. Faure. *Simulation du mouvement Brownien et des diffusions*. PhD thesis, Ecole Nationale des Ponts et Chaussées, 1992.
- [7] M. B. Giles. Improved multilevel Monte Carlo convergence using the Milstein scheme. In *Monte Carlo and quasi-Monte Carlo methods 2006*, pages 343–358. Springer, Berlin, 2008.
- [8] M. B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.
- [9] M. B. Giles, D. J. Higham, and X. Mao. Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff. *Finance Stoch.*, 13(3):403–413, 2009.
- [10] P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. Probability and Mathematical Statistics.
- [11] S. Heinrich. Monte Carlo complexity of global solution of integral equations. *J. Complexity*, 14(2):151–175, 1998.
- [12] S. Heinrich. Multilevel monte carlo methods. *Lecture Notes in Computer Science, Springer-Verlag*, 2179(1):58–67, 2001.
- [13] S. Heinrich and E. Sindambiwe. Monte carlo complexity of parametric integration. *J. Complexity*, 15(3):317–341, 1999. Dagstuhl Seminar on Algorithms and Complexity for Continuous Problems (1998).
- [14] J. Jacod. On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de Probabilités, XXXI*, volume 1655 of *Lecture Notes in Math.*, pages 232–246. Springer, Berlin, 1997.
- [15] J. Jacod and P. Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26(1):267–307, 1998.
- [16] A. Kebaier. Statistical romberg extrapolation: A new variance reduction method and applications to option pricing. *Annals of Applied probability*, 15(4):2681–2705, 2005.
- [17] P. E. Kloeden and E. Platen. Numerical methods for stochastic differential equations. In *Nonlinear dynamics and stochastic mechanics*, CRC Math. Model. Ser., pages 437–461. CRC, Boca Raton, FL, 1995.
- [18] R. Korn, E. Korn, and G. Kroisandt. *Monte Carlo methods and models in finance and insurance*. Chapman & Hall/CRC Financial Mathematics Series. CRC Press, Boca Raton, FL, 2010.
- [19] T. G. Kurtz and P. Protter. Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. In *Stochastic analysis*, pages 331–346. Academic Press, Boston, MA, 1991.
- [20] T. G. Kurtz and P. Protter. Weak error estimates for simulation schemes for SDEs. 1999.

- [21] B. Lapeyre and E. Temam. Competitive Monte carlo methods for the pricing of Asian options. *Journal of Computational Finance*, 5(1), 2001.
- [22] P. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990.
- [23] A. Rényi. On stable sequences of events. *Sankhyā Ser. A*, 25:293–302, 1963.
- [24] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.